



Continuous-time Quantum Optimisation

by

Robert James Banks

Supervisors:

Paul A. Warburton

&

Dan E. Browne

A thesis submitted in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy

Department of Physics and Astronomy
University College London

Tuesday 1st October, 2024

Declaration of Authorship

I, Robert BANKS, confirm that the work presented in this thesis, titled “Continuous-time quantum optimisation”, is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. The following parts of this thesis have been published in journals or in preparation:

1. **Chapters 2, 7 & Appendix B:** ROBERT J. BANKS, DAN E. BROWNE AND P. A. Warburton, *Rapid quantum approaches for combinatorial optimisation inspired by optimal state-transfer*, Quantum, 8 (2024), p. 1253, [doi:10.22331/q-2024-02-13-1253](https://doi.org/10.22331/q-2024-02-13-1253).
2. **Chapters 2, 3, 4, 5 & 6:** ROBERT J. BANKS ET AL., *Continuous-time quantum walks for MAX-CUT are hot*, Quantum, 8 (2024), p. 1254, [doi:10.22331/q-2024-02-13-1254](https://doi.org/10.22331/q-2024-02-13-1254).
3. **Chapters 2, 3, 6 & Appendix A:** ROBERT J. BANKS, GEORGIOS S. RAFTIS, DAN E. BROWNE AND P. A. Warburton, *Continuous-time quantum optimisation without the adiabatic principle*, arXiv preprint, (2024), 2407.03910, [url:arxiv.org/abs/2407.03910](https://arxiv.org/abs/2407.03910).

Abstract

The adiabatic theorem presents a clear bottleneck on adiabatic quantum optimisation. Even given access to a coherent quantum system, the time required to remain adiabatic is typically too long to be reached when solving a large combinatorial optimisation problem. This necessitates operating the device non-adiabatically. Continuous-time quantum walks, multi-stage quantum walks, and reverse-quantum annealing all present attempts to use the same hardware, while dropping the adiabatic requirement. Since these approaches operate far from adiabaticity, the adiabatic theorem cannot be used to motivate their use in tackling combinatorial optimisation problems.

Continuous-time quantum walks have been shown to perform well numerically on some optimisation problems. However, the mechanism behind quantum walks for optimisation has not been well understood. By establishing a connection between continuous-time quantum walks and the eigenstate thermalisation hypothesis, this dissertation explores the mechanism behind continuous-time quantum walks as well as how they can be optimised.

By appealing to pure-state statistical physics more generally, it is shown how a variety of time-dependent approaches, such as multi-stage quantum walks, can be motivated. This is done by using the physically motivated assumption, termed Planck's Principle, that work cannot be extracted from a cyclic process in an isolated system. This is sometimes referred to as Kelvin's formulation of the second law of thermodynamics.

ABSTRACT

This work also explores a different design mantra, away from adiabatic inspired approaches, based on optimal state transfer. This provides insight on how continuous-time quantum algorithms might be designed away from conventional approaches. It is shown that the optimal state transfer approaches can outperform current conventional quantum approaches.

Impact Statement

Many problems can be framed as finding the maximum or the minimum from a finite number of choices, these are called combinatorial optimisation problems. Examples of combinatorial optimisation include: scheduling deliveries to minimise the distance travelled or buying stocks to maximise profit. The number of possible choices can be very large, for example, there are 9×10^{157} possible routes to try for 100 deliveries. Searching over them all, to find the shortest one is not possible. The large number of possible solutions makes solving combinatorial optimisation problems very hard. Yet combinatorial optimisation problems are very common and providing effective approaches to solve them will have a high impact in a range of disciplines.

It has been proposed that quantum technologies might help to solve combinatorial optimisation problems, by exploiting novel quantum effects. Designing quantum algorithms is difficult, especially without mature hardware to test ideas on. Quantum algorithms can be discrete-time (typically implemented by gates) or continuous-time, where the quantum algorithm is carried out on an analogue quantum system. Many approaches for solving combinatorial optimisation problems in analogue quantum systems have relied on the adiabatic theorem. Essentially, by varying parameters in the quantum system slowly, it is possible to find the solution to the combinatorial optimisation problem. In practice, going sufficiently slowly to reach the desired solution is not possible. Therefore, recent approaches have explored going faster within this framework. Numerical studies have been performed to establish the

performance of these approaches, but the motivation and understanding behind them has been largely missing.

The impact of this work is to provide intuition behind these faster continuous-time quantum approaches for combinatorial optimisation problems. This is done by establishing links between these approaches and statistical physics. In the near term, this shows how some continuous-time quantum optimisation algorithms should be implemented. In the longer term, better understanding should allow for better design of continuous-time quantum optimisation approaches. This dissertation also explores an alternative design framework, far from the adiabatic theorem, based on optimal state transfer, opening up an alternative route for algorithm design.

Outside of quantum algorithm design, this dissertation presents a step towards understanding the potential impact of quantum algorithms for optimisation. Given the ubiquity of these problems, any speed up or increased accuracy may find use within industry.

Acknowledgements

This dissertation would not have been possible without the support of many people. First amongst them Professor Paul Warburton. Thanks to Paul, I have taken my first steps into research and begun to understand what a sentence is. It has been a pleasure to have been part of the supportive environment he has fostered over the years. Thank you for being a great supervisor and giving me a chance.

At times, I have struggled with understanding my own research. Many thanks to Professor Dan Browne for co-supervising me, whose patience and insightful probing questions have made this dissertation what it is.

This dissertation could not have been achieved without Dr Abbie Bray's support. Thank you for introducing me to Orbyts, and all the wonderful members, teachers, and students. Carrying out research with school students was a fulfilling experience, that I will not be forgetting anytime soon. Thank you for this dissertation template, being a mentor, and most of all a wonderful friend.

My PhD experience would not have been the same without peers to share in it with. Many thanks to my fellow research group members. Thank you to Natasha Feinstein, my fellow theorist, for putting up with my ramblings (and calling my nonsense out when needed). Leon Guerrero, you made the office a better place to be, I still count myself lucky that you joined our research group, thank you. Thank you to (honorary group member) Emily Cook for the superb company, laughs, and cheesecake. Thank you to Henry Chew for your patience, I cannot imagine a better person to sit next to in the office, hopefully we can finish that

ACKNOWLEDGEMENTS

chess game soon. Pias Tubsrinuan, thank you for lighting up the office. George Raftis, it was an absolute pleasure having you in the LCN, thank you for keeping me on my toes and your company. You have all been excellent colleagues and even better friends.

My UCL experience started in my bedroom during the Covid pandemic. Many thanks to Cohort 7 who have been with me from the start. To Jack Dunham, Lesley Grover, Evan Sutcliffe, Ana Villanueva Ruiz De Temino, and Tim Weaving, thank you for keeping me smiling from the start.

None of this would have been possible without the support of my colleagues at the LCN, including the professional services teams. Thank you for making the LCN a desirable place to work.

To Matt Cooper, thank you for helping me get on the cluster and all your support both practical and emotional. Thank you for being a very good friend.

Thank you to Frankie Griffin, my housemate over several years, who always had time and food for me. Thank you to all my housemates, who have given me a space to work, relax and call home.

Finally, I would not be writing this without the constant love and support from my friends and family. To my grandparents who have always been willing to help-out and fostering a love for learning, one bicarbonate volcano at a time, thank you. To my parents who packed me off to school dressed as a Dalek or with an electromagnet in my school bag, thank you for raising a nerd. To my siblings, Matthew and Lucy, thank you for your infectious laughs, smiles and keeping me humble. This dissertation is dedicated to all of you.

Contents

Declaration of Authorship	i
Abstract	iii
Impact Statement	v
Acknowledgements	vii
List of Figures	xiii
List of Tables	xxvii
List of Acronyms	xxix
1 Introduction	1
2 Combinatorial optimisation and continuous-time quantum optimisation approaches	5
2.1 A primer on notation and terminology	5
2.2 Combinatorial optimisation	6
2.3 Encoding combinatorial optimisation problems	7
2.4 Limitations of quantum computing	13
2.5 Choice of metrics	14
2.6 Continuous-time quantum optimisation approaches	16
2.6.1 Adiabatic quantum optimisation	17
2.6.2 Quantum annealing	17
2.6.3 Continuous-time quantum walks	19

2.6.4	Multi-stage quantum walks	21
2.6.5	Reverse quantum annealing	22
2.7	The Quantum Approximate Optimisation Algorithm . . .	23
2.7.1	Separating discrete time quantum algorithms from continuous-time quantum algorithms	25
2.8	Summary	28
2.9	A note on numerical work and presentation	30
3	Pure-state statistical physics	33
3.1	Introduction	33
3.2	Conservation of Energy	34
3.3	Eigenstate Thermalisation Hypothesis	34
3.4	Spin-flip symmetry and the ETH	38
3.5	The second-law of thermodynamics in an isolated-system	41
3.5.1	Diagonal Entropy	41
3.5.2	Planck's Principle	42
3.6	Conclusion	46
4	Continuous-time quantum walks for Max-cut are hot	47
4.1	Introduction	47
4.2	Optimising the free parameter	48
4.3	Continuous-time quantum walks are well modelled by thermal states	52
4.4	Modelling the density of states	55
4.5	Predictions from the density of states	61
4.6	Estimating $\overline{\langle H_p \rangle}$	65
4.7	The performance of CTQWs for Max-cut	70
4.8	Conclusion	71
5	Thermalisation timescales	73
5.1	Introduction	73
5.2	A geometrical description of the dynamics	73
5.3	The very short-time limit	79
5.4	An Ansatz for the Schrödinger Evolution	82

5.5	A thermal start	88
5.6	Conclusion	89
6	Statistical-mechanics applied to continuous-time quantum optimisation	91
6.1	Introduction	91
6.2	Multi-stage quantum walks	92
6.2.1	Motivating multi-stage quantum walks	92
6.2.2	A thermal model for multi-stage quantum walks	94
6.2.3	How many stages?	97
6.2.4	Consequences of passivity	101
6.2.5	Introduction of a bath	104
6.3	Pure State Thermal Quantum Annealing	106
6.3.1	Pure state thermal quantum annealing is path-independent	107
6.3.2	Diagonal entropy is conserved	108
6.3.3	Numerical evidence of PSTQA	109
6.3.4	Ansatz based approach	112
6.4	Cyclic approaches	113
6.4.1	Warm starting continuous-time quantum walks does not work (on average)	113
6.4.2	Reverse quantum annealing	114
6.4.3	Biased quantum annealing	117
6.5	Conclusion	121
7	Rapid Hamiltonians for combinatorial optimisation problems	123
7.1	Hamiltonian design	124
7.1.1	The optimal Hamiltonian	124
7.1.2	Adapting the optimal Hamiltonian	127
7.2	Taking the commutator between the initial and final Hamiltonian	129
7.2.1	The optimal approach for a single qubit	129
7.2.2	Application to larger problems	132

7.2.3	Direct numerical comparisons to QAOA $p = 1$. . .	150
7.3	An improvement inspired by the Quantum Zermelo problem	152
7.3.1	The approach	152
7.3.2	Numerical simulations	155
7.4	Using knowledge of the initial state	157
7.5	Conclusion	163
8	Conclusion	167
8.1	Summary	167
8.2	Outlook on continuous-time quantum optimisation	168
8.2.1	Accepting and exploiting dynamics	170
8.2.2	Incorporating more information	171
8.2.3	Shifting the focus away from ground-state probabilities	172
A	Further numerical work on CTQO	175
A.1	Numerical validation of PSTQA	175
A.2	Ansatz approaches towards PSTQA	180
A.2.1	The Gaussian model	180
A.2.2	The Gaussian model applied to Max-cut	183
A.2.3	The exponentially modified Gaussian model	183
A.2.4	The exponentially modified Gaussian model applied to Max-cut	185
A.3	Warm-starting CTQWs	186
A.4	Numerically observing cyclic cooling	190
B	Further numerical work on rapid Hamiltonians for optimisation	199
B.1	A further numerical study on a Sherrington-Kirkpatrick inspired model	199
B.2	Instances for which QAOA $p = 1$ outperforms H_1	202
	Bibliography	207

List of Figures

2.1	A cartoon of a typical MSQW schedule. The ellipsis indicates the possibility of more (or fewer) stages.	22
2.2	A gate based implementation of a ZZ-rotation.	26
2.3	An illustration of a box-plot for a data set, named y . The minimum value of y is denoted by $\min(y)$, the maximum value is denoted by $\max(y)$	31
3.1	The solid black line sketches the typical density-of-states $\Omega(E)$, for a large non-integrable system [1]. The solid blue line shows a possible energy distribution for the initial state. The red line shows a possible resulting energy distribution as the result of a cyclic process. The energy distribution has been broadened and has moved towards the middle of the spectrum.	43
4.1	Plots of possible metrics for assessment of CTQWs for the randomly generated 12-qubit instance, shown in the inset of Fig. 4.1a	49
4.2	The time-averaged value $\langle H_p \rangle$ as γ is changed for the 12 qubit graph shown in the inset of Fig. 4.1a. The dashed purple line shows the location of γ_{opt} and $\overline{\langle H_p \rangle}_{\text{min}}$	49
4.3	A comparison of the optimised value of γ , i.e., γ_{opt} and the heuristic choice of γ shown in Eq. 4.4 for 100 binomial graphs at each problem size. The number of qubits n is equal to the number of vertices in the Max-cut graph.	51

4.4	A comparison of the optimised value of γ , i.e., γ_{opt} and the heuristic choice of γ shown in Eq. 4.4 for three-regular graphs. In ascending problem size, there are: 16, 45 problem instances.	52
4.5	The performance of a CTQW with γ optimised to give the best value of $\overline{\langle H_p \rangle}$. The dashed purple line shows the thermal state prediction. The temperature is fixed using Eq. 4.5.	53
4.6	The inverse temperature associated with CTQWs (i.e. the solution to Eq. 4.5). For each problem instance γ is equal to γ_{opt} , as shown in Fig. 4.3a and Fig. 4.4a.	54
4.7	The difference between the true value of $\overline{\langle H_p \rangle}$ and the thermal prediction (for the values of β shown in Fig. 4.6) normalised by the magnitude of the true value of $\overline{\langle H_p \rangle}$ for each problem.	54
4.8	Histogram for the DOS for two graphs with optimised γ . The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution.	58
4.9	Histogram for the DOS for two graphs with γ optimised. The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution. The dashed red line shows the EMG distribution.	61
4.10	The error in the predicted inverse temperature assuming a Gaussian DOS (i.e. Eq. 4.21). For each problem instance γ is equal to γ_{opt} , shown in Fig. 4.3a and Fig. 4.4a.	62
4.11	The error in the predicted inverse temperature assuming an EMG DOS (i.e. Eq. 4.22). For each problem instance γ is equal to γ_{opt} , shown in Fig. 4.3a and Fig. 4.4a.	63

4.12	The entanglement entropy averaged over a hundred 14-qubit binomial instances (solid blue line). The dashed purple line shows $S^{\text{norm}}/2$, the dashed green line $S^{\text{EMG}}/2$, and the dotted red line shows $S^{\text{SM}}/2$. The shaded regions show a single standard deviation.	65
4.13	The solid blue line shows $\overline{\langle H_p \rangle}$ for a single instance of a 12-qubit binomial graph. The dashed purple (pink) line shows the prediction from Eq. 4.26 assuming a Gaussian (EMG) DOS.	68
4.14	The difference in $\overline{\langle H_p \rangle}$ between the optimal choice of γ and the prediction from the EMG DOS (γ_{EM}), normalised by the absolute value of $\overline{\langle H_p \rangle}$	69
4.15	The difference between the prediction of $\langle H_p \rangle$ with an EMG DOS using Eq. 4.26 (i.e. $\overline{\langle H_p \rangle}_{\text{EM}}$) and the true value. For each instance, the value of γ has been chosen such that $\overline{\langle H_p \rangle}_{\text{EM}}$ is minimised. Each bar shows one hundred binomial graph instances.	69
4.16	Estimated scaling of the CTQW for binomial graphs, assuming thermalisation and an EMG DOS. Each data point shows the median performance of one hundred binomial graph instances.	70
4.17	The normalised performance in terms of the time averaged mean cut value (i.e., $(\kappa_2 - \overline{\langle H_p \rangle})/2$) of the CTQW for the optimal γ predicted by the EMG DOS compared to the maximum cut (i.e. $(\kappa_2 - E_0^{(p)})/2$) on the graph	71

5.1	The performance of a CTQW on two different 12-qubit binomial graphs with different values of γ . The blue line shows the result of direct integration of the Schrödinger equation. The pink line shows the result of the two-dimensional approximation (i.e, Eq. 5.22) The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$. The dashed green line in (5.1b) shows the location of $n\mathcal{T}^{-\frac{1}{4}}$, with $n = 12$. The values of γ have been chosen to illustrate different behaviour and do not correspond to the optimal γ for the graph.	81
5.2	The values of $\nu_{m,n}$ for an 11-qubit binomial graph with $\gamma = 1$	83
5.3	An example of a CTQW on a 12-qubit binomial graph. The solid blue line shows the Schrödinger evolution of $\langle H_p(t) \rangle$. The solid red line shows Eq. 5.35 with τ given by Eq. 5.34. The dashed green line shows the location of τ . The dotted purple line shows the location of τ extracted numerically from the Schrödinger evolution (i.e. Eq. 5.36).	84
5.4	The thermalisation timescale for a closed-system CTQW on n qubits. For each instance, 100 examples are considered.	85
5.5	The error between the Gaussian Ansatz and the Schrödinger equation (i.e. Eq. 5.41) at $t = \mathcal{T}^{-1/4}$. At each problem size, 100 instances at Max-cut on binomial graphs are considered with $\gamma = \sqrt{n/\kappa_2}$. The value of $\langle H_p \rangle$ has been estimated using an EMG Ansatz for the DOS.	87
5.6	An example of a CTQW on a 10-qubit binomial graph with $\gamma_1 = 0.705$. The solid blue line shows the Schrödinger evolution of $\langle H_p(t) \rangle$. The solid red line shows Eq. 5.44 with τ given by Eq. 5.46. The dashed green line shows the location of τ . The dotted purple line shows the location of τ extracted numerically from the Schrödinger evolution (i.e. Eq. 5.36). The initial state is a Gibbs state with $\beta = 0.5$ and $\gamma_0 = 0.440$. .	87

5.7	The thermalisation timescale for a closed-system CTQW on n qubits. For each instance, 100 examples are considered. The initial state is a Gibbs state with $\beta = 0.5$ and $\gamma_0 = (3/4)\sqrt{(n/\kappa_2)}$. For each instance $\gamma_1 = (6/5)\sqrt{(n/\kappa_2)}$	89
6.1	The cyclic schedule used to demonstrate the performance of MSQWs.	92
6.2	A five-stage CTQW on a 12-qubit binomial graph. The purple line is the numerical prediction, assuming that the MSQW is well modelled by a CTQW. The pink line shows the prediction from Eq. 4.26 and modelling the DOS as an EMG distribution.	95
6.3	The difference between the stationary state calculated from the Schrödinger equation and a thermal state for five-stage CTQW on 100 12-qubit binomial graphs. “EMG DOS” refers to the prediction with the energy fixed by Eq. 6.6. The numerical approach gives the prediction from using the numerically determined energy, numerically fixing the temperature and calculating $\langle H_p \rangle$	96
6.4	The monotonic schedules used to demonstrate that the use of more stages results in a lower final energy.	97
6.5	The schedules used to demonstrate that more stages result in a lower value of $\langle H_p \rangle$	99
6.6	The solid blue line sketches a possible $\gamma(t)$, which in this case is monotonically increasing. The purple line shows the piece-wise constant function used to approximate $\gamma(t)$. The dashed pink line shows the cyclic process used to demonstrate $\langle H_p(t) \rangle \leq \langle H_p(0) \rangle$	102
6.7	The 10-qubit Max-cut example.	109
6.8	The schedule used for the 10-qubit Max-cut instance. The pink line shows $A(t)$ and the blue line shows $B(t)$	110

6.9	A comparison of $\langle H_p(t) \rangle$ calculated from the Schrödinger equation (blue) and the PSTQA equations (pink) for the 10 qubit example. The dashed purple line shows the ground-state energy, $E_0^{(p)}$, of the problem Hamiltonian.	110
6.10	The diagonal entropy calculated from the Schrödinger equation for the 10 qubit Max-cut example.	110
6.11	The inverse temperature calculated from the PSTQA equations for the 10 qubit Max-cut example.	111
6.12	PSTQA for the 10-qubit Max-cut graph shown in Fig. 6.7. The blue line shows the Schrödinger evolution. The green line shows the prediction from the PSTQA equations when an exponentially modified Gaussian ansatz is made for the density-of-states.	112
6.13	The effect on the problem Hamiltonian of introducing H_b . The initial state is highlighted in purple. States lower in energy are coloured blue, while states higher in energy are coloured pink. As long as $ \Delta W < \alpha$, the expectation value of $\langle H_p \rangle$ will decrease.	118
7.1	The geometric intuition behind finding the Hamiltonian for optimally transferring between the ground-states of H_d and H_p on the Bloch sphere. The vectors $\pm \hat{m}$ ($\pm \hat{n}$) are the eigenvectors of H_d (H_p). The aim is to generate a rotation of θ around \hat{k} to map $\pm \hat{m}$ to $\pm \hat{n}$. The handedness of the cross-product takes into account the direction.	130
7.2	A time-domain plot of the ground-state probability (in pink, corresponding to the right-hand side y-axis) and approximation ratio (in blue, corresponding to the left-hand side y-axis) for H_1 applied to Max-cut on a 2-regular graph with 400 qubits. Random guessing corresponds to a ground-state probability of $2^{-399} \approx 10^{-120}$. The dashed purple line shows the location of the optimal time, corresponding to the maximum in approximation ratio.	139

7.3	Performance of H_1 on Max-cut with 2-regular graphs. Only even numbers of qubits are plotted.	139
7.4	The optimal times for Max-cut on two-regular graphs. The blue (pink) line shows the time that optimises the approximation ratio (ground state probability). Even qubit numbers only.	140
7.5	The performance of H_1 on Max-cut with two-regular graphs on an odd number of qubits	140
7.6	A cartoon for calculating the Lieb-Robinson Bound (LRB) on Max-cut with two-regular graphs. Each node represents a qubit, and each edge the interactions between them. To estimate a local expectation value, say $Z_j Z_{j+1}$, the local subgraph in pink is simulated. Bounding the error requires ΔH , corresponding to all the interactions leaving the subgraph. In this case, this corresponds to all the interactions which connect the blue qubits to the pink ones.	143
7.7	LRB applied to Max-cut on two-regular graphs. The dashed blue line shows the local estimate of $\langle Z_j Z_{j+1} \rangle$ and the shaded violet region the allowed region from the LRB. The red line shows the actual value for 400 qubits. The dynamics of H_1 closely resembles the local simulation.	144
7.8	The Lieb-Robinson inspired bound for subgraph 3 in Table 7.1. This subgraph dominates the worst-case bound. The LRB has a minimum at around a time of 0.1 while the locally estimated value has a minimum at around a time of 0.2. . . .	145
7.9	The performance of H_1 on randomly generated instances of three-regular graphs. For each problem size, 100 instances were generated. After accounting for graph isomorphisms, the number of samples in order of ascending problem size were [15, 46, 87, 97]. Disconnected graphs were allowed. . . .	146

7.10 Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots) for three-regular graphs. The dashed purple line shows QAOA and H_1 performing the same.	147
7.11 Performance of H_1 on 100 randomly-generated instances of Max-cut.	148
7.12 Example subgraphs	149
7.13 Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots) for the problem instances considered in Sec. 7.2.2. Fig. 7.13a compares approximation ratios. Fig. 7.13b compares the optimal times (i.e. the time that maximises the approximation ratios) of the two approaches. The dashed purple line corresponds to equal performance.	151
7.14 A cartoon of the evolution of states in the QZ problem for constant H_{QW} . In the interaction picture, with background Hamiltonian H_{QW} , it appears the final state is moving under the influence of this Hamiltonian. In this frame, Eq. 7.1 can then be applied. It then remains to move out of the interaction picture to get Eq. 7.52.	153
7.15 The performance of Eq. 7.55 on two-regular graphs. The legend shows the order of T , with ‘exp’ referring to Eq. 7.54. The dashed lines show the asymptotic performance of QAOA. . .	155
7.16 The performance of the QZ-inspired approach on 100 random Max-cut instances. The x-axis label refers to the order of T in the expansion of Eq. 7.55, with 0 being H_1 and e referring to the full exponential (i.e. Eq. 7.54)	157
A.1 The time evolution of $\langle H_p(t) \rangle$ for a Max-cut instance on a 10-qubit binomial graph. The schedule is a linear ramp, with a minimum value of 0.3 and maximum value 1.3. The blue line shows the Schrödinger evolution. The pink line shows the solution of the PSTQA equations for this instance.	175

A.2	The inverse temperature for the Max-cut instance considered in Fig. A.1 from the PSTQA equations.	176
A.3	The diagonal entropy for the Max-cut instance considered in Fig. A.1 from the Schrödinger equations.	176
A.4	A box-plot showing $(\langle H_p(t_f) \rangle - \langle H_p(t_f) \rangle_{TQA}) / \langle H_p(t_f) \rangle$ for Max-cut on 10 qubit binomial graphs. The final value of the Schrödinger evolution is denoted by $\langle H_p(t_f) \rangle$. The value predicted by the PSTQA equations is denoted by $\langle H_p(t_f) \rangle_{TQA}$. In each case, a linear schedule is used with $A(0) = B(t_f) = 1.3$ and $A(t_f) = B(0) = 0.3$. At each value of t_f , 90 instances are considered.	176
A.5	A box-plot showing the change in diagonal entropy for Max-cut on 10 qubit binomial graphs for the instances shown in Fig. A.4.	177
A.6	The schedule used for the 10-qubit SKM instance	178
A.7	A comparison of $\langle H_p(t) \rangle$ calculated from the Schrödinger equation (the solid blue line) and the PSTQA equations (the solid pink line) for the 10 qubit SKM instance. The schedule is shown in Fig. A.6	178
A.8	The diagonal entropy calculated from the Schrödinger equation for the 10 qubit SKM example (Fig. A.7).	179
A.9	The inverse temperature calculated from the PSTQA equations for the 10 qubit SKM example (Fig. A.7).	179
A.10	A box-plot showing $(\langle H_p(t_f) \rangle - \langle H_p(t_f) \rangle_{TQA}) / \langle H_p(t_f) \rangle$ for 10 qubit SKM instances. The final value of the Schrödinger evolution is denoted by $\langle H_p(t_f) \rangle$. The value predicted by the PSTQA equations is denoted by $\langle H_p(t_f) \rangle_{TQA}$. In each case a linear schedule is used with $A(0) = B(t_f) = 1.1$ and $A(t_f) = B(0) = 0.1$. At each value of t_f 90 instances are considered.	179
A.11	A box-plot showing the change in diagonal entropy of the 10-qubit SKM instances shown in Fig. A.4.	180

A.12 The evolution of $\langle H_p(t) \rangle$ for a Max-cut instance on 13 qubit binomial graph. The schedule is linear with $A(0) = B(t_f = 10) = 1.1$ and $A(t_f = 10) = B(0) = 0.1$. The blue line shows the Schrödinger evolution, and the pink line the prediction using an exponentially modified Gaussian ansatz for the density-of-states.	185
A.13 The error between $\langle H_p(t_f) \rangle$ from the Schrödinger equation and the exponentially modified Gaussian model, $\langle H_p(t_f) \rangle_{EMG}$. The figure shows 100 Max-cut instances on 13 qubit binomial graphs. The schedule is linear with $A(0) = B(t_f = 10) = 1.1$ and $A(t_f = 10) = B(0) = 0.1$. The prediction from the exponentially modified Gaussian is denoted by $\langle H_p(t_f) \rangle_{EMG}$, and the true value by $\langle H_p(t_f) \rangle$. The x-axis shows $E_{(0)}^p$, the ground state energy of H_p	186
A.14 The time-evolution of $\langle H_p \rangle$ for a 12 qubit instance. The dashed purple line shows the infinite time average, $\overline{\langle H_p \rangle}$. . .	187
A.15 The time-averaged distribution of $\langle H_p \rangle$ for the problem instances shown in Fig. A.14. The dashed purple line shows the original value of $\langle H_p \rangle$. The dashed pink line shows the average of the distribution.	187
A.16 The time-averaged value of $\langle H_p \rangle$ as g is varied for the Max-cut instance considered in Fig. A.14.	188
A.17 The initial value of $\langle H_p \rangle$ for a warm-started CTQW compared to the infinite time averaged value. For each instance $g = 0.5$. The blue circles show 100 Max-cut instances. The pink diamonds show 100 SKM instances. The number of qubits is 12 in all cases. The initial state is randomly selected given that it satisfies Eq. 6.53. The dashed purple line is the "y = x" line for ease of visual comparison.	188

A.18	The initial value of $\langle H_p \rangle$ for a warm-started CTQW compared to the infinite time averaged value. For each instance $g = 0.5$. The blue circles show 100 Max-cut instances. The pink diamonds show 100 SKM instances. The number of qubits is 12 in all cases. The initial state is given by Eq. A.28.	189
A.19	The schedule $G(t)$ appended to the driver Hamiltonian, used for the cyclic processes described in Appendix A.4.	192
A.20	RQA for a 10 node Max-cut graph. The pink dots show $\langle H_p \rangle$ for the post-selected distribution. The blue diamonds show $\langle H_p \rangle$ as sampled from application of U_{cyc}	192
A.21	The diagonal entropy S_d for RQA applied to a 10-node Max-cut graph. The pink dots show S_d for the post-selected distribution. The blue diamonds show S_d as sampled after application of U_{cyc} . The logarithm used to calculate the diagonal entropy is taken to be base 2.	192
A.22	The inverse probability (Eq. A.33) for each stage of RQA on a 10-vertex Max-cut graph.	193
A.23	The performance of QA (purple circles), RQA (blue diamonds) and BQA (pink crosses) for a 12 qubit example. . . .	194
A.24	The extractable work, W , for the 12-qubit instances considered in Fig. A.23. The blue diamonds correspond to RQA. The pink crosses BQA. The purple circles show W for BQA neglecting the change in energy of $H_b^{(l)}$, i.e, the change in $\langle H_p \rangle$	194
A.25	The average approximation ratio for the 12 qubit instances at each shot number. Since the approaches terminate at different shot numbers, the number of instances decreases. The pink crosses show the average approximation ratio for BQA. The blue diamonds show the average approximation ratio for RQA. The lines show one standard deviation. The dashed purple line shows an approximation ratio of 1. The final decrease in approximation ratio in BQA reflects more difficult problems for BQA, since it has not terminated after many shots.	195

A.26	The extractable work for the 12 qubit instances at each shot number. Since the approaches terminate at different shot numbers, the number of instances decreases. The pink crosses show the average extractable work for the BQA protocol. The blue diamonds show the average extractable work for RQA. The lines show one standard deviation. The purple circles show the change in $\langle H_p \rangle$ for BQA.	195
B.1	Performance of H_1 on 100 randomly-generated instances of SKM.	200
B.2	Width of the final distribution, σ for randomly generated instances of Max-cut and SKM.	200
B.3	Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots). The dashed purple line corresponds to equal performance.	201
B.4	Performance on the QZ-inspired approach on 100 instances of the SKM. The x-axis label refers to the order of T in the expansion of Eq. 7.55, with 0 being H_1 and e referring to the full exponential (i.e. Eq. 7.54).	201
B.5	The performance of H_1 compared to QAOA $p = 1$ on Max-cut with three-regular graphs. The instances with atypical run-times for H_1 are highlighted in pink. The rest of the data has been faded for clarity and to show how the remaining data is clustered, with darker regions corresponding to more data points. The dashed purple line shows both approaches performing the same.	203

-
- B.6 The performance of H_1 compared to QAOA $p = 1$ on Max-cut with three-regular graphs. The highlighted pink dots show the instances with atypical H_1 run-times, optimised to give the best possible approximation ratio with shorter run-times than QAOA $p = 1$. The new run-times are shown in the lower plot, with the updated approximation ratio plotted in the upper plot. The dashed purple line shows both approaches performing the same. 204
- B.7 The performance of H_1 compared to QAOA $p = 1$ on Max-cut with randomly generated graphs. The instances with atypical run-times for H_1 are highlighted in blue. The rest of the data has been faded for clarity and to show how the remaining data is clustered, with darker regions corresponding to more data points. The dashed purple line shows both approaches performing the same. 205
- B.8 The performance of H_1 compared to QAOA $p = 1$ on Max-cut with randomly generated graphs. The instances with atypical run-times for H_1 are highlighted in blue, these have been optimised to give the best possible approximation ratio with a corresponding run-time smaller than the QAOA $p = 1$ optimal time. The new run-times are shown in the lower plot, with the updated approximation ratio plotted in the upper plot. The rest of the data has been faded for clarity. The dashed purple line shows both approaches performing the same. 205
- B.9 The performance of H_1 compared to QAOA $p = 1$ on the SKM. Here we neglect smaller problem instances, plotting problem sizes between 6 and 12 qubits. 206

List of Tables

2.1	Summary of CTQO approaches and QAOA. The table summarises the initial state and the schedules $A(t)$ and $B(t)$, including the boundary conditions at the initial time $t = 0$ and the final time $t = T$. The asterisk * denotes the required addition of a third term that biases towards the initial state. . . .	29
7.1	Numerical details for the LRB applied to H_1 on Max-cut with three-regular graphs. Each column shows, from top to bottom, the local subgraph being simulated; the local estimate of $Z_i Z_j$ to be minimised; the corresponding worst case from the LRB; and the corresponding cut value for this worst case. All of these values are taken at the optimised time of 0.093. . .	145
B.1	The exceptions for the three-regular graphs.	202

List of Acronyms

ABC = Anti-periodic Boundary Conditions
AQO = Adiabatic Quantum Optimisation
BQA = Biased Quantum Annealing
CTQW = Continuous-Time Quantum Walk
DOS = Density-Of-States
EMG = Exponentially Modified Gaussian
ETH = Eigenstate Thermalisation Hypothesis
IQR=Interquartile Range
LRB = Lieb-Robinson Bound
MSQW = Multi-Stage Quantum Walk
NISQ = Near-term Intermediate-Scale Quantum
PBC = Periodic Boundary Conditions
PSTQA = Pure-State Thermal Quantum Annealing
QA = Quantum Annealing
QAOA = Quantum Approximate Optimisation Algorithm *or* Quantum Alternating Operator Ansatz
QZ = Quantum Zermelo
RQA = Reverse Quantum Annealing
SKM = Sherrington-Kirkpatrick inspired Model

Chapter 1

Introduction

Combinatorial optimisation problems have been proposed as a possible candidate for useful application of a quantum computer [2, 3, 4]. Perhaps the first evidence for this came with Grover’s eponymous algorithm [5]. The problem Grover was attempting to solve was the unstructured search problem: how to find a marked element in a list of length N , given an oracle to validate the solution. This is a search problem, which classically requires on the order N calls to the oracle. Grover showed that this could be done with order \sqrt{N} calls to the oracle on a quantum computer, a quadratic speed-up. This speed-up has been shown to be optimal for this problem [6]. Grover’s algorithm can be adapted to combinatorial optimisation problems [7, 8].

Grover’s algorithm treats the combinatorial optimisation problem as a black-box oracle that can be queried. Ultimately, constructing this oracle requires knowledge of the combinatorial optimisation problem [9]. Grover’s algorithm sets a lower bound on what might be achieved by a quantum device, where no information about the combinatorial optimisation problem is being utilised. It was recently questioned if Grover’s algorithm constitutes quantum speed-up at all [9]. This motivates the need for other quantum algorithms for combinatorial optimisation problems.

Kadowaki and Nishimori introduced quantum annealing in [2], the idea being that the quantum algorithm would have direct access to the cost function associated with the combinatorial optimisation problem. Quan-

tum fluctuations could then be used to search and escape from local minima in the cost function. The approach was largely framed as a heuristic approach. Adiabatic quantum optimisation, proposed by Farhi *et al.* [10] demonstrated that combinatorial optimisation problems could be solved exactly through adiabatic evolution. A superpolynomial speed up using an oracle model with adiabatic quantum computing has been demonstrated by Hastings [11]. An exponential speed-up has also been demonstrated on the oracular glued-trees problem [12]. However, this speed up requires transitions to and from the first excited state. The exponential speed-up also vanishes if the classical computer can exploit noise in the system [13].

Both adiabatic quantum optimisation and quantum annealing are designed to be implemented on an analogue quantum system, hence they are referred to as continuous-time quantum algorithms. Broadly speaking, quantum annealing is considered an example of a near-term intermediate-scale quantum (NISQ) algorithm [4]. NISQ generally refers to quantum algorithms that may provide some form of advantage in the near-term without the need for full error correction. Quantum annealing and adiabatic quantum optimisation helped to inspire the gate-based quantum approximate optimisation algorithm [14]. Further background material on combinatorial optimisation problems and continuous-time quantum optimisation algorithms can be found in Chapter 2.

Analogue quantum systems have been engineered in an attempt to realise quantum annealing, termed quantum annealers [15, 16, 17]. These devices have limited coherence times, and it is well known that adiabatic quantum optimisation requires exponentially long run times [18, 19, 20, 21]. Even given the perfect quantum annealer, these run-times could not be realised. This has warranted further exploration of continuous-time quantum approaches that involve shorter runtimes on the quantum device, by trading off short run-times for performance or a larger sampling cost. These approaches are more amenable for NISQ devices. A candidate continuous-time quantum algorithm that trades off

shorter individual runtimes for a higher sampling cost is the continuous-time quantum walk [22]. Continuous-time quantum walks involve a time-independent Hamiltonian run for a short time and sampled. These have been shown to scale better than Grover’s algorithm numerically for certain combinatorial optimisation problems [22]. Continuous-time quantum walks have a free parameter that needs to be chosen to optimise the performance. In Chapter 4 the long-time behaviour of a continuous-time quantum walk is explored. The work in this chapter is motivated by the eigenstate thermalisation hypothesis [23, 24] that attempts to explain how thermalisation occurs in a closed quantum system. By assuming thermalisation, it is shown how the free-parameter in a continuous-time quantum walk can be chosen.

When implementing a continuous-time quantum walk, it is necessary to choose a sampling time. This is explored in Chapter 5 by looking at geometrical properties, such as the torsion of the wave function [25]. This is then built on by using an Ansatz for the time evolution. Much of the work on CTQWs makes use of sensible Ansätze to convert a problem of large matrices to a problem with continuous variables, parameterised by a few real numbers.

The natural extension of a continuous-time quantum walk is referred to as multi-stage quantum walk. The performance of this approach remains largely unexplored, as well as the mechanism behind this approach. In Chapter 6 pure-state statistical physics, introduced in Chapter 3, is used to provide motivation for this approach, namely, the physically motivated assumption that work cannot be extracted from a cyclic process in an isolated system. This is sometimes referred to as Planck’s Principle [26, 27] or Kelvin’s formulation of the second law of thermodynamics [1]. Since Planck’s principle is naturally formulated in terms of cyclic processes, cyclic quantum annealing is also investigated in Chapter 6.

All the approaches so far have some common features. All are arguably quantum annealing with a specific choice of Ansatz (and maybe a dif-

ferent initial state). Perhaps a fair criticism of the field is that selecting a new Ansatz is sufficient to claim a novel approach. This is not a surprise, as all the approaches stem from quantum annealing and are often motivated in relation to the adiabatic theorem. Designing novel approaches outside the adiabatic inspired framework is not straightforward, especially without mature and flexible quantum hardware. In Chapter 7, an approach inspired by Hamiltonians for optimal state-transfer is proposed. This is compared to conventional quantum approaches with similar resources.

Chapter 2

Combinatorial optimisation and continuous-time quantum optimisation approaches

This chapter outlines the background material needed to understand the quantum algorithms discussed in this dissertation. The first half of this chapter will give a very brief introduction to combinatorial optimisation problems, including some relevant examples. The second half will introduce some established continuous-time quantum optimisation (CTQO) algorithms designed to tackle combinatorial optimisation problems.

2.1 A primer on notation and terminology

This dissertation adopts the standard convention of setting \hbar equal to one. The Boltzmann constant k_B is also set to one throughout this dissertation. The respective Pauli matrices are denoted by X , Y and Z . The identity is denoted by I . Angular brackets, for example, $\langle A \rangle$, denote the expectation value of an observable (in this case, the expectation of the observable A). The state, for which the expectation value should be taken with respect to, should be clear from context.

Some of the algorithms discussed in this work do not need to be performed on a (universal) quantum computer. Many CTQO approaches are currently performed on specialised devices with limited coherence

[15, 16, 28, 29]. To avoid any possible contention, we refer to devices on which the algorithms are run on as quantum processing units (QPUs) [15].

The QPUs are assumed to be qubit based. Certain implementations of CTQO approaches require the addition of extra qubits. Here we refer to these qubits as auxiliary qubits.

2.2 Combinatorial optimisation

An instance of a combinatorial optimisation problem consists of two parts: a finite set of feasible solutions \mathcal{F} and a cost-function C [30]. The aim is to find an element x in \mathcal{F} that satisfies:

$$C(x) \leq C(y) \text{ for all } y \text{ in } \mathcal{F}. \quad (2.1)$$

The element x is said to be a global minimum. Here, the optimisation problem has been framed in terms of finding a minimum. Since maximisation and minimisation are equivalent under a sign-flip of the cost function (i.e. minimising C is equivalent to maximising $-C$) maximisation is accounted for as well in this formulation. Shifting the cost-function by a constant, or multiplication by a positive number, leaves the ordering between elements of \mathcal{F} the same. A combinatorial optimisation problem is a collection of instances [30].

Let us consider an example of an instance of a combinatorial optimisation problem, where an asset manager is looking to select 13 assets from a possible 25. The set of feasible solutions is all possible combinations of 13 assets, a grand total of 5,200,300 combinations. The cost function is the expected loss on a combination of assets. The global optimum is the combination of assets that makes the smallest loss (or equivalently the greatest profit). Enumerating 5,200,300 possible combinations from 25 assets is daunting. As the problem size is increased by an order of magnitude, say to selecting 130 assets from 250, there are now 7.5×10^{73}

combinations to try. Trying all possible combinations is definitely not a viable strategy. It is this scaling and size of the set of feasible solutions that makes combinatorial optimisation problems hard to solve. In practice the set \mathcal{F} will not be all possible configurations since the asset manager will have to meet certain constraints – for example the risk from a combination of stocks should not exceed a given value. Combinatorial optimisation problems are very common, with application in finance [31, 32, 33], logistics [34, 35], machine learning [36, 37], and biology [38, 39].

2.3 Encoding combinatorial optimisation problems

Before discussing in depth the encoding of combinatorial optimisation problems, it is useful to outline the entire workflow:

1. The combinatorial optimisation problem is encoded and passed to a QPU.
2. The QPU then undergoes a run, of which there might be many.
3. At the end of each run, the qubits are sampled. The result of sampling the QPU is a string, s , consisting of 1s and 0s.
4. A decoding procedure, \mathcal{D} , is then applied to this string to find a possible solution to the optimisation problem, $\mathcal{D}(s) = x$.

At the end of each run, the cost of x according to the classical cost function can be calculated by $C(x) = C(\mathcal{D}(s))$. So the composition of $C(\cdot)$ and $\mathcal{D}(\cdot)$ can be considered a cost function on the string s , denoted by C_s . Typically, it is expected C_s will take the form (up to some irrelevant constant):

$$C_s(s) = \sum_i h'_i s_i + \sum_{i,j} J'_{i,j} s_i s_j + \sum_{i,j,k} J'_{i,j,k} s_i s_j s_k + \sum_{i,j,k,l} J'_{i,j,k,l} s_i s_j s_k s_l + \dots, \quad (2.2)$$

2. COMBINATORIAL OPTIMISATION AND CONTINUOUS-TIME QUANTUM OPTIMISATION APPROACHES

where s_i is the value of the i^{th} element in s , corresponding to the i^{th} qubit. The coefficients $h'_i, J'_{i,j}, J'_{i,j,k}, J'_{i,j,k,l} \dots$ are real numbers depending on the problem instance. The ellipsis in Eq. 2.2 denotes the presence of higher order terms. In the context of quantum mechanics, the string s is best understood as a state vector, denoted by $|s\rangle$, where:

$$C_s(s) = C'_s(|s\rangle), \quad (2.3)$$

up to operations that preserve the ordering on the feasible strings. Rewriting the cost function in terms of matrix multiplication gives:

$$C'_s(|s\rangle) = \langle s | \hat{C}_s | s \rangle, \quad (2.4)$$

where \hat{C}_s is a Hermitian operator. The operator \hat{C}_s can be found by replacing s_i in Eq. 2.2 with $(I - Z_i)/2$ where Z_i is the Pauli Z operator acting on the i^{th} qubit. The result is:

$$\hat{C}_s = \sum_i h_i Z_i + \sum_{i,j} J_{i,j} Z_i Z_j + \sum_{i,j,k} J_{i,j,k} Z_i Z_j Z_k + \sum_{i,j,k,l} J_{i,j,k,l} Z_i Z_j Z_k Z_l + \dots \quad (2.5)$$

Since \hat{C}_s corresponds to a Hermitian operator acting on the qubits, it is easily identified as corresponding to a Hamiltonian. In this context, the optimisation problem becomes finding the ground-state of \hat{C}_s . It is the coefficients (i.e. $h_i, J_{i,j} \dots$) in \hat{C}_s that are typically fed into the QPU.

The original cost-function $C(x) = C(\mathcal{D}(s))$ was defined only on x that are elements of the feasible subspace. Not all strings s (or equivalently $|s\rangle$) will correspond to a feasible x . This could be down to constraints in the combinatorial optimisation problem or limitations of the QPU. For example, in an analogue quantum system where the Hamiltonian \hat{C}_s is physically realised, higher order couplings are typically difficult to engineer [40]. Therefore, auxiliary qubits are typically added such that

\hat{C}_s is quadratic and takes the form,

$$\hat{C}^{\text{Ising}} = \sum_i h_i Z_i + \sum_{i,j} J_{i,j} Z_i Z_j, \quad (2.6)$$

this is known as an Ising Hamiltonian. Even then, connectivity between qubits might be limited, requiring further auxiliary qubits. The introduction of the auxiliary qubits or other constraints from the optimisation problem will result in strings s that are not in \mathcal{F} . The constraints are enforced by making sure that the quantum algorithm does not, or is very unlikely to, output a string that does not correspond to an unfeasible solution. This is typically done by adding large energy penalties to the cost function [41], so unfeasible strings are very unfavourable. Or alternatively, by designing the quantum algorithm so it only searches feasible solutions [42, 43]. The formulation of the cost function \hat{C}_s , or more typically \hat{C}^{Ising} , plus imposing constraints, constitutes encoding the combinatorial optimisation problem. It remains then to determine what $\mathcal{D}(\cdot)$ is. This will depend on the nature of x ; for example, if x is a binary variable [41], an integer [44], or a graph [45].

In this dissertation, the feasible solution space \mathcal{F} for the combinatorial optimisation problems considered is the set of all possible spin configurations. The problems considered are naturally formulated as Ising Hamiltonians, and it is assumed a fully connected architecture is available. This minimises the encoding and decoding required, so that the focus is on the quantum algorithm, while maximising their possible implementation in analogue quantum systems without specifying a hardware model. Since the problems have no constraints, only \hat{C}^{Ising} needs to be formulated. The rest of this section details the encoding of the combinatorial optimisation problems considered in this dissertation.

The Ising model

The Ising model is the focus of almost all quantum algorithms for combinatorial optimisation problems [10, 14]. The feasible set of solutions

2. COMBINATORIAL OPTIMISATION AND CONTINUOUS-TIME QUANTUM OPTIMISATION APPROACHES

for the Ising model is the set of all possible configurations of n spins, where n is a positive integer. Each spin s can either be spin up, in which case $s = 1$, or spin down such that $s = -1$. There are 2^n possible configurations of spins. Denoting a single spin configuration by $(s_0, s_1, \dots, s_{n-1})^T$, the cost function is given by

$$C(\vec{S})^{\text{Ising}} = \sum_{i>j} J_{i,j} s_i s_j + \sum_i h_i s_i, \quad (2.7)$$

where the couplings $J_{i,j}$ and fields h_i are real parameters that need to be set. The associated Hamiltonian is the Ising Hamiltonian:

$$H_p^{(\text{Ising})} = \sum_{i>j} J_{i,j} Z_i Z_j + \sum_i h_i Z_i, \quad (2.8)$$

where the notation has been changed to make clear that this is a Hamiltonian. All other combinatorial optimisation problems considered in this dissertation correspond to specific choices of the h_i 's and $J_{i,j}$'s in Eq. 2.8.

Max-cut

Given some graph, $G = (V, E)$ the aim of Max-cut is to find the maximum cut. A cut separates the nodes in a graph G into two disjoint sets. The length of the cut is equal to the number of edges between the two disjoint sets. The maximum-cut is the cut with the greatest length. Given a graph with n vertices, there are 2^{n-1} possible partitions to try. Hence, the number of possible cuts grows exponentially with the problem size. For a Max-cut instance, the feasible set of solutions is all possible bipartitions of the graph. The cost function that is to be minimised is

$$-1 \times (\text{length of the cut}). \quad (2.9)$$

Max-cut can easily be expressed in terms of spin variables. A cut splits a graph into two disjoint sets, here denoted by **0** and **1**. Let the spin variable $s_k \in \{-1, 1\}$ denote which set the k^{th} node is in, such that if

$s_k = 1$, the node is contained in the set $\mathbf{0}$. If $s_k = -1$ then the node is contained in $\mathbf{1}$. The length of the cut is the number of edges between the two sets and is given by

$$C_{\max} = \frac{1}{2} \sum_{(i,j) \in E} 1 - s_i s_j, \quad (2.10)$$

since only anti-aligned spins contribute. Since, we are primarily interested in minimisation, the cost function is given by

$$C_{\min} = \frac{1}{2} \sum_{(i,j) \in E} s_i s_j - 1. \quad (2.11)$$

For simplicity, as the ordering between cuts is not changed by a constant shift or scaling by a positive factor of the cost function, often the cost function is taken to be:

$$C_{\text{simp}} = \sum_{(i,j) \in E} s_i s_j. \quad (2.12)$$

Identifying the spin variables with Pauli matrices gives:

$$H_p^{(MC)} = \sum_{(i,j) \in E} Z_i Z_j, \quad (2.13)$$

again the notation has been changed to make clear that this is a Hamiltonian. Max-cut has a trivial representation as an Ising model.

Max-cut is defined on a graph. In this work, we primarily focus on three types of graph.

1. Two-regular graphs (i.e. a ring). This is sometimes referred to as an anti-ferromagnetic ring or the ring of disagrees. This is a trivial problem.
2. Three-regular graphs. A graph where each node is connected to exactly three others. These graphs are sometimes referred to as cubic graphs.

3. Binomial graphs. Binomial graphs are generated by selecting an edge between any two nodes with probability p . Max-cut undergoes a computational phase change for random graphs at $p = 1/2$ (harder problems appear for $p > 1/2$) [76–78]. We set $p = 2/3$ for this dissertation (note that this is no guarantee of hardness for the problem instances considered). These graphs are sometimes referred to as Erdős-Rényi graphs.

This dissertation focuses on Max-cut. This is a standard choice for exploring quantum algorithms for combinatorial optimisation [14, 46] due to its simple form.

Sherrington-Kirkpatrick inspired model

In a Max-cut problem, all the couplers in the graph are set to the same value. Here we introduce a second problem, the Sherrington-Kirkpatrick inspired model (SKM) [22, 47], where this is not the case. The problem is to find the ground-state of

$$H_p^{(SK)} = \sum_{i,j} J_{i,j} Z_i Z_j \quad (2.14)$$

where the $J_{i,j}$'s are randomly selected from a normal distribution with mean 0 and variance 1 [47]. Each qubit is coupled to every other qubit.

To further distinguish the SKM from the Max-cut problems, a bias term can be introduced to the Hamiltonian,

$$H_p^{(SKM)} = \sum_{i,j} J_{i,j} Z_i Z_j + \sum_i^n h_i Z_i, \quad (2.15)$$

where the h_i 's are also randomly selected from a normal distribution with mean 0 and variance 1. We use the SKM in the Appendices to give further numerical evidence of the claims in this dissertation away from Max-cut.

2.4 Limitations of quantum computing

Given that the set of feasible solutions \mathcal{F} is typically too large for problems of interest to be enumerated, better approaches need to be found. Before discussing quantum approaches, it is worth understanding what might be the limitations on the resources required to find the global minimum of a set \mathcal{F} .

Complexity theory is outside the scope of this dissertation. It does however contain some pertinent lessons to understand the limitations of quantum computers. The following exposition on complexity theory is based on [48] and [49]. Complexity theory attempts to sort problems into classes depending on how hard they are to solve. Informally, **NP** consists of decision problems that are hard to solve on a classical computer but whose solution is easy to verify. A decision problem is a question whose answer is ‘Yes’ or ‘No’. A computation is hard if it requires an exponential amount of time to solve. A computation is easy if the time required in a classical computer is polynomial. Easy decision problems are said to belong to **P**. Consider the decision problem, closely related to Max-cut: does there exist a cut of length d or more for a given graph? This is a decision problem, that becomes more challenging as the graph size is increased.

Max-cut is an example of an **NP** problem. This is a statement about the worst case; there are however easy Max-cut examples. If a problem is at least as hard as every problem in **NP**, then it is **NP-hard**. This means any problem in **NP** can be efficiently converted to a **NP-hard** problem. If a problem is **NP-hard** and in **NP**, it is said to be **NP-complete**. In general, Max-cut is an **NP-complete** problem [50]. Indeed, finding very good approximations for Max-cut is a computationally hard problem [51]. Finding the ground-state of an Ising Hamiltonian is an **NP-complete** problem as well. From a complexity theory perspective, studying an **NP-complete** problem allows for the study of any other **NP** problem. It is generally believed that some problems are in **NP** but not in **P**, i.e.

there exist some hard problems.

The class of problems that can be solved in polynomial time on a quantum computer is called **BQP**. It is generally believed that not every problem in **NP** is in **BQP**. This means that it is expected that at least some optimisation problems will require an exponential runtime on a quantum computer. So what advantage might quantum computers present for classical combinatorial optimisation problems? For certain optimisation problems, quantum computing might provide an exponential speed up [12] or a sub-exponential speed-up [11]. In some cases, the speed-up might come from a prefactor advantage [52] or better scaling in the exponent [53].

2.5 Choice of metrics

Assessing the performance of an algorithm is challenging. A common choice is the ground-state probability, P_{gs} . Calculating the ground-state probability requires knowing the ground-state, which is not known for problems of interest, and to warrant the use of a QPU it should be hard to find the ground-state in the first place. Another choice of metric that combines both run time and ground state probability is the time to solution [48, 54]. This makes for fairer comparison for quantum algorithms with very different runtimes. Again, this relies on knowing the ground-state probability.

The ground-state probability is a reasonable measure when the only solutions of interest are global minimums. But, it fails to capture the performance of an approximate solver (a solver that finds *good enough* solutions). A common measure for approximate solvers is the approximation ratio. The approximation ratio is a measure on the final distribution produced by the approach. The approximation ratio is $\langle H_p \rangle / E_0^{(p)}$ where $E_0^{(p)}$ is the energy of the ground-state solution and, H_p is the problem Hamiltonian encoding the optimisation problem. The expectation value is taken with respect to the final state. If the approach finds the

ground state exactly, then $\langle H_p \rangle / E_0^{(p)} = 1$. All the problem Hamiltonians considered in this dissertation do not consist of terms proportional to the identity (i.e. H_p is chosen to be traceless). This is to simplify the analysis, while preserving the ordering between eigenstates of H_p . The result is that the approximation ratio can be positive or negative. Random guessing (for traceless H_p) has an approximation ratio of 0. A negative approximation ratio means that the approach is doing worse than random guessing on average. This choice is made to achieve consistency across different problems. This is distinct from papers that include terms proportional to the identity in the Hamiltonian. For example, in [14] they use terms proportional to the identity, meaning that random guessing achieves a non-zero approximation-ratio.

The approximation ratio has some major flaws: it requires knowledge of $E_0^{(p)}$ which is typically difficult to find, it makes no reference to time, and it makes no mention of the width of the distribution. A metric close to time to solution was proposed for approximate solvers in [53], though again this required detailed knowledge of the combinatorial optimisation problem. An alternative metric is $\langle H_p \rangle$. This preserves the ordering between approaches and approximation ratio and does not require knowledge of $E_0^{(p)}$, but it is not normalised. Such a quantity would be verifiable on a quantum computer, but still inherits the other drawbacks from the approximation ratio. Since $\langle H_p \rangle$ is not normalised, it is not straightforward to compare between different problem instances or problem sizes, given the possible variation in energy scale. This makes it difficult to extrapolate the performance.

There is unlikely to be a single metric that will capture all the desired characteristics of a quantum algorithm. In this dissertation, the focus is on $\langle H_p \rangle$ as it captures approximate solvers as well as corresponding to a sum over local observables that could feasibly be measured in a physical system.

2.6 Continuous-time quantum optimisation approaches

This section outlines some popular CTQO algorithms at a high-level. The aim of CTQO is to find or approximate solutions to a combinatorial optimisation problem using a quantum device such as a quantum computer or a quantum annealer. All the CTQO approaches discussed in this chapter contain some common features. For each approach, the combinatorial optimisation problem is encoded as the energy levels of a Hamiltonian H_p , which is typically an Ising Hamiltonian, with lower energy eigenstates corresponding to better solutions. There is a second Hamiltonian, referred to as a driver Hamiltonian, denoted by H_d . The typical requirements on H_d are that it does not commute with H_p , and its ground-state is known and easy to prepare. The conventional choice is the transverse-field Hamiltonian given by

$$H_d^{(TF)} = - \sum_i X_i, \quad (2.16)$$

where X_i is the Pauli X matrix acting on the i^{th} qubit. This is the common choice in CTQO due to its straight-forward implementation. The ground state of $H_d^{(TF)}$ is denoted by $|+\rangle$. In CTQO the system is evolved under the Hamiltonian

$$H_{\text{CTQO}} = A(t)H_d + B(t)H_p, \quad (2.17)$$

or slightly modified versions of this Hamiltonian. The distinction between many CTQO approaches is in the choice of the initial state, and the schedules, denoted by $A(t)$ and $B(t)$. The total time of the evolution is denoted by T . We refer to this set-up as the CTQO framework.

2.6.1 Adiabatic quantum optimisation

Adiabatic quantum optimisation (AQO) initialises the system in the ground state of H_d . The system is evolved for a total time T , with the boundary conditions $A(0) = B(T) = 1$ and $A(T) = B(0) = 0$. The schedules are chosen such that the evolution is adiabatic, meaning that the system remains in its instantaneous ground-state at all times. The result is a quantum algorithm that transfers the system from the known ground state of H_d to the unknown ground state of H_p , exactly solving the optimisation problem.

This relies on the system being isolated, resulting in unitary evolution. The other major requirement is for the system to remain adiabatic, or equivalently, that the evolution must be sufficiently slow such that no transitions to higher energy levels occurs. AQO has been shown to achieve the same scaling as Grover's algorithm on the unstructured search problem [55]. For many problems, the time required to remain adiabatic is exponentially increasing with the problem size [18, 19, 20, 21, 48]. Due to this exponentially increasing time, various other approaches have been proposed, such as: modifying the adiabatic path (through counter diabatic driving [56, 57, 58, 59, 60]).

2.6.2 Quantum annealing

Quantum annealing (QA) is a heuristic version of AQO. QA shares the same initial state and boundary conditions on the schedules as AQO. The schedules are typically monotonic too. The adiabatic requirement is however typically dropped. QA sometimes includes the incorporation of open quantum system effects [61, 62]. The explicit choice of causing excitations in QA is sometimes referred to as diabatic quantum annealing [12, 63].

QA was first proposed in analogy with simulated annealing [2]. In simulated annealing, the algorithm mimics thermal fluctuations to search the solution space and escape local minima. By analogy, QA uses quan-

tum fluctuations to search the solution space and escape local minima by quantum tunnelling. Since the probability of quantum tunnelling decays exponentially with the width of the barrier, it has been proposed that QA should work well in energy landscapes with tall and thin barriers [64].

Max-cut on two-regular graphs is a well studied problem in the context of QA [10]. The problem Hamiltonian for a two-regular graph is given by

$$H_p = \sum_{i=1}^n Z_i Z_{i+1} \quad (2.18)$$

with $n + 1 = 1$, corresponding to periodic boundary conditions. The Hamiltonian consists of nearest-neighbour terms only. The performance of QA on this problem has been understood by applying the Jordan-Wigner transformation to map the problem onto free fermions [10]. They found that AQO could solve Max-cut on two regular graphs in a time that scales as n^3 .

Braida *et al.* [46] have investigated QA on three-regular graphs with short run times. By treating the algorithm as a local algorithm and using Lieb-Robinson inspired bounds, they were able to lower-bound the performance of QA on three regular graphs. This approach does not necessarily find a tight bound, but it is nonetheless impressive, with bounds in continuous-time quantum computation a rarity. They showed that QA finds at least 0.5933 times the best cut. This was later improved to 0.7020 without monotonic schedules and a tighter bound [65].

A number of hardware platforms have been proposed for realising CTQO. Two of the leading platforms are superconducting qubits [16, 28, 53] and Rydberg atoms [17, 66]. Rydberg atoms, due to their intrinsic physics, are well suited to solving a combinatorial optimisation called the maximum-independent set. D-Wave, a commercial quantum technology company, currently manufactures annealers featuring over 5000 superconducting qubits [28] and has demonstrated coherent evolution for times on the order of tens of nanoseconds [16, 28].

Munoz Bauza and Lidar have claimed to have demonstrated a scaling advantage over known classical algorithms for approximating a certain combinatorial optimisation problem on a D-Wave device [53]. This scaling advantage is based off a problem native to the architecture of the D-Wave and sampling from a specific set of couplings. They made a comparison to a single classical algorithm, that has previously been demonstrated to scale better than QA. A polynomial scaling was observed in the quantum and classical case, with the quantum annealer providing a better scaling, provided the desired level of approximation was not too high.

2.6.3 Continuous-time quantum walks

Continuous-time quantum walks (CTQWs) for optimisation are a relatively new approach [22]. Just as in QA and AQO, the system is initialised in the ground state of the driver Hamiltonian. The schedule is time-independent such that $A(t) = 1$ and $B(t) = \gamma$, where γ is a real number that needs to be set. The CTQW Hamiltonian is then given by

$$H_{\text{CTQW}} = H_d + \gamma H_p. \quad (2.19)$$

When H_d is the transverse-field Hamiltonian, the CTQW corresponds to a quantum walk on the Boolean hypercube with each node weighted by the classical cost-function ¹.

The quantities that need to be set are the runtime T of the CTQW and γ . The performance of CTQWs has been observed not to be strongly dependent on T for combinatorial optimisation problems, provided T is sufficiently large [22, 68]. The main challenge in realising CTQWs, from

¹Note that in previous works, the parameter γ has been appended to the driver Hamiltonian [22, 67]. Here it has been appended to the problem Hamiltonian for ease of analysis. This is because there are two important energy scales in a CTQW: the energy of the initial state and the energy of the Hamiltonian. By appending γ to the problem Hamiltonian the energy of the Hamiltonian can be tuned independently of the energy associated with the initial state. Further to this, later in the dissertation we will explore $\langle H_p \rangle$ by varying γ , so the free parameter is appended to H_p for consistency.

2. COMBINATORIAL OPTIMISATION AND CONTINUOUS-TIME QUANTUM OPTIMISATION APPROACHES

an algorithmic perspective, is setting the free parameter γ to optimise the performance of the CTQW. If the absolute value of γ is too small, the system approximately evolves under H_d , resulting in trivial dynamics. If the absolute value of γ is too large, then the system evolves approximately under γH_p , again resulting in trivial dynamics. The optimal γ must sit within these two limits.

The system is initialised in the ground state of the driver Hamiltonian, so the energy associated with the driver (i.e. $\langle H_d(t) \rangle$) can only increase compared to its initial value. The system is time-independent, so energy is conserved. Therefore, $\gamma \langle H_p(t) \rangle$ can only decrease from its initial value. Provided $\gamma > 0$, the energy associated with H_p can only decrease. The initial-state corresponds to random guessing, so CTQWs can never do worse than random guessing on average, with $\gamma > 0$. If $\gamma < 0$ then the CTQW would never do better than random guessing on average. In conclusion, γ must be greater than 0 but not so large that the Hamiltonian is well approximated by γH_p .

CTQWs were first examined in the context of solving SKM instances by Callison *et al.* [22]. The choice of γ was selected on the mantra of maximising dynamics, achieved by matching the energy scales of the driver and problem Hamiltonians. They found that the ground-state probability scaled better than Grover's for the problems considered. The mantra of maximising dynamics was further explored in [67] and extended to other combinatorial optimisation problems. CTQWs have also been shown to recover the same scaling as Grover's algorithm for the unstructured search problem [69].

The physical motivation behind CTQWs for optimisation has been largely unexplored. In Chapters 4 and 5, we explore the intuition behind this approach and show how this can lead to physically motivated choices of γ as well as a timescale.

2.6.4 Multi-stage quantum walks

Multi-stage quantum walks (MSQWs) are an extension of CTQWs with close analogy to QA. The initial state is the ground-state of the driver Hamiltonian. The schedule associated with the driver Hamiltonian is constant in time. The schedule associated with the problem Hamiltonian is a piece-wise constant non-decreasing function. Each interval where the Hamiltonian is held constant is referred to as a stage. The Hamiltonian is given by

$$H_{MSQW} = H_d + \Gamma(t)H_p, \quad (2.20)$$

where, the schedule $\Gamma(t)$ is given by

$$\Gamma(t) = \begin{cases} \gamma_1 & \text{for } 0 \leq t \leq t_1 \\ \gamma_2 & \text{for } t_1 < t \leq t_2 \\ \dots & \\ \gamma_k & \text{for } t_{k-1} < t \leq t_k \\ \dots & \\ \gamma_l & \text{for } t_{l-1} < t, \end{cases} \quad (2.21)$$

where $\gamma_1 < \gamma_2 < \dots < \gamma_k < \gamma_l$ and $0 < t_1 < t_2 \dots t_k \dots t_l$. Fig. 2.1 shows a cartoon of a typical MSQW schedule. It has been numerically observed that the expectation of H_p decreases (corresponding to better solutions on average) as $\Gamma(t)$ is increased [67]. It is insufficient to appeal to the adiabatic theorem to explain the performance, and there are limited theoretical motivations for this approach. From energy conservation mechanisms, it has been shown that this approach can always do better than random guessing [67]. As this protocol is piece-wise constant with the time dependence entering through sudden quenches, the approach is more amenable to analytic investigation. The performance of MSQWs remains largely unexplored.

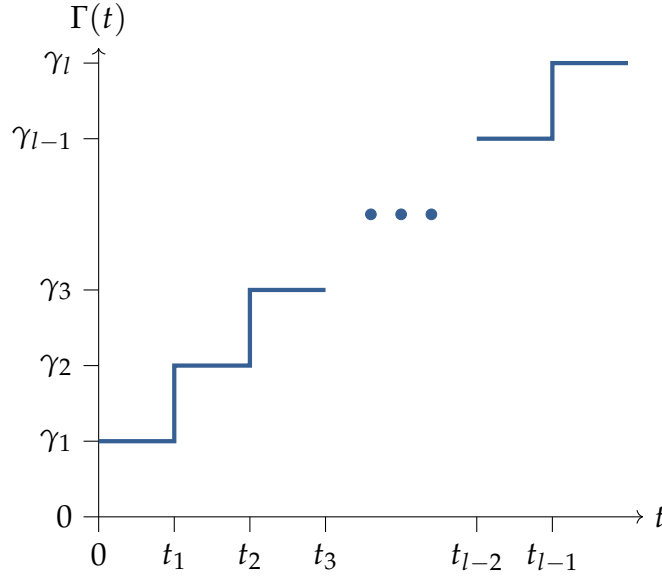


Figure 2.1: A cartoon of a typical MSQW schedule. The ellipsis indicates the possibility of more (or fewer) stages.

2.6.5 Reverse quantum annealing

Before attempting to solve a combinatorial optimisation problem, it is reasonable to assume that a classical approach might have been tried first, resulting in a candidate string or access to a distribution of candidate strings that can be sampled. Or that given a previous run of a quantum annealer, one now has access to a candidate solution. Given that for a typical problem the quantum annealer might be operating non-adiabatically, one expects to be sampling numerous times from the quantum annealer. Each conventional run of a quantum annealer makes no use of any prior information or candidate solutions that might be available. Reverse quantum annealing (RQA) [70, 71] is an attempt to make use of known candidate solutions. RQA can be implemented on current noisy hardware [72].

RQA is accomplished by operating the schedules cyclically. The system is initialised in an eigenstate of the problem Hamiltonian, with the problem Hamiltonian turned on. This is then turned down, and the driver Hamiltonian turned on. The driver Hamiltonian is then turned

off, and the system returns to H_p . The relative strength of the driver to the problem Hamiltonian controls how locally the system searches [70, 72]. Then there is a classical update step to decide how the string should be initialised in the next anneal [70, 71].

Other variants on RQA include biasing the anneal with a third term in the Hamiltonian. In which case the Hamiltonian is given by

$$H_{\text{BQA}} = A(t)H_d + B(t)H_p + C(t)H_b, \quad (2.22)$$

where the initial state is an eigenstate of H_p and the ground state of H_b . Both H_p and H_b are diagonal in the computational basis. The schedule $A(t)$ is cyclic and satisfies $A(0) = A(T) = 0$. The initial and final Hamiltonian can be some combination of H_p and H_b . Since sampling only occurs in the computational basis and the initial state is diagonal in the computational basis, the schedules can always be chosen such that $B(0) = B(T)$ and $C(0) = C(T)$. To distinguish between RQA with a bias from the version without, we refer to RQA with a bias as biased quantum annealing (BQA). This distinction is made since the addition of a bias changes the computational mechanism, as discussed in Sec. 6.4.3.

BQA was explored in [73] in the context of solving spin-glass problems and implemented on a D-Wave quantum annealer in [74]. It was argued that the biasing can help drive the system from the spin glass phase, for improved searching.

2.7 The Quantum Approximate Optimisation Algorithm

The quantum approximate optimisation algorithm (QAOA) [14] is a gate-based design philosophy for determining the schedules, $A(t)$ and $B(t)$. The initial state is the ground state of the driver Hamiltonian. In QAOA $A(t) = 1 - B(t)$, and $B(t) = 0$ or $B(t) = 1$. The switching parameter, p , controls the number of times the schedules alternate be-

2. COMBINATORIAL OPTIMISATION AND CONTINUOUS-TIME QUANTUM OPTIMISATION APPROACHES

tween the two parameter settings. The duration between switching is either determined prior to the evolution by a classical computer or by using a classical variational outer-loop [14, 75], attempting to minimise $\langle H_p \rangle$ by varying $2p$ free-parameters. In short, the quantum computer is used to prepare the Ansatz:

$$|\psi_{\text{QAOA}}\rangle = \prod_{k=1}^p e^{-iH_d\beta_k} e^{-iH_p\gamma_k} |+\rangle, \quad (2.23)$$

where the parameters $\beta_1, \beta_2, \dots, \beta_p, \gamma_1, \gamma_2, \dots, \gamma_p$ have been optimised so that $|\psi_{\text{QAOA}}\rangle$ minimises $\langle \psi_{\text{QAOA}} | H_p | \psi_{\text{QAOA}} \rangle$.

QAOA again relies on the adiabatic principle to provide a guarantee of finding the ground state in the limit of infinite p . However, far from this limit, the variational method allows QAOA to exploit non-adiabatic evolution [76]. QAOA has also become a popular choice for benchmarking the performance of quantum hardware [77, 78, 79].

QAOA applied to Max-cut on two-regular graphs can be studied using the Jordan Wigner transformation [80, 81]. Alternatively, provided p is sufficiently small, the performance of QAOA can be understood in terms of locality [14] for two-regular graphs. Due to the structure of the Ansatz in QAOA (shown in Eq. 2.23), to find the expectation of a term in H_p , such as $Z_i Z_{i+1}$, it is only necessary to consider a subgraph. This subgraph consists of all nodes connected by no more than p edges to a node in the support of the expectation value being calculated. For two regular graphs this is a chain consisting of $2p + 2$ nodes. Provided this subgraph is smaller than the problem graph (i.e. the two-regular graph has more than $2p + 2$ nodes), then QAOA is operating locally. For a given p , all the subgraphs are identical for a two-regular graph. Hence, the performance of QAOA for this problem depends only on its performance on this subgraph. As a direct consequence of the locality, the approximation ratio of QAOA will not change as the size of the two-regular graph is scaled. By optimising over this subgraph with a classical resource, it is therefore possible to find the optimal time and

approximation ratio of QAOA for this problem.

Max-cut on three-regular graphs was considered in the original QAOA paper by Farhi *et al.* [14]. The local-nature of QAOA allowed them to calculate explicit bounds on the performance of their algorithm. To this end, the graph was broken down into subgraphs to measure local expectation values (i.e., $\langle Z_i Z_{i+1} \rangle$). For QAOA $p = 1$, there are three distinct subgraphs. By simulating QAOA for the subgraphs and bounding the proportion of subgraphs in the problem, they calculated a lower bound on the performance. The small number of relevant subgraphs for three-regular graphs makes this approach particularly amenable. They found that QAOA $p = 1$ will produce a distribution whose average will correspond to at least 0.6924 times the best cut. This lower-bound is saturated by triangle-free graphs. These graphs consist of a single subgraph. Therefore, the performance of QAOA $p = 1$ is largely dependent on the proportion of edges in this problem that belong to this subgraph.

2.7.1 Separating discrete time quantum algorithms from continuous-time quantum algorithms

Is QAOA a continuous-time quantum algorithm or a discrete-time quantum algorithm? To better explain the difference between discrete (typically gate-based) and continuous-time quantum algorithms, consider the example of making a ZZ -rotation of an angle θ , i.e.,

$$R_{ZZ}(\theta) = e^{-i\frac{\theta}{2}ZZ}. \quad (2.24)$$

In a continuous-time quantum algorithm, a ZZ interaction would be engineered between the two qubits. The system would then be evolved for a time $T = \theta/2$ and the correct rotation generated. Importantly, if the evolution was paused halfway through, the rotation generated would still be a ZZ rotation.

Conversely, in a gate-based system, there is a set of hardware native gates which the ZZ -rotation is decomposed into. Consider a native

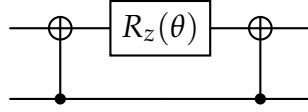


Figure 2.2: A gate based implementation of a ZZ-rotation.

gate set of single-qubit gates and Controlled-NOT gates, then the ZZ-rotation might be achieved as shown in Fig. 2.2 [82]. If this evolution was paused halfway through, then the rotation generated would not be a ZZ-rotation.

In other words, in a continuous-time quantum algorithm the Hamiltonian description is a true description at all times. In a discrete-time quantum computer, the Hamiltonian description need only be accurate at the end of each gate, i.e, at a discrete number of time points. Discrete quantum algorithms can allow for fault-tolerant error correction [82]. There is no fault-tolerant model of error correction for continuous-time quantum algorithms.

This raises the question of why one should study continuous-time quantum algorithms. Currently, scaleable fault-tolerant quantum computing has not been achieved [83], leaving scope to explore more near-term quantum technologies to gain intuition and to search for possible advantages [4]. Quantum annealing, an example of a CTQO algorithm, has been proposed as a possible candidate for near-term advantage [4]. The analogue nature of continuous-time quantum algorithms bring them closer to experiment, though with some limited programmability. It has been argued that the analogue nature leads to some CTQO algorithms having a degree of inbuilt error resilience [84, 85, 86].

Further to the above, CTQO encourages the design of quantum algorithms in terms of Hamiltonians. This can help provide intuition into quantum algorithms. Indeed, QAOA was inspired by QA/AQO [14], which are undeniably CTQO algorithms. Ultimately, CTQO approaches might one day be mapped onto discrete quantum hardware, detaching

from the analogue hardware [87].

Finally, we return to the question of whether QAOA is a continuous-time quantum algorithm. The answer is not clear cut. Consider the case where the driver Hamiltonian is the transverse field and H_p is a quadratic Ising Hamiltonian. If the native gate-set of the quantum computer consists of ZZ-rotations and single qubit rotations, and no SWAP-gates are required, then QAOA is a continuous-time quantum algorithm. If finite-connectivity results in the need for SWAP gates, then there are two choices:

1. Like QA, auxiliary qubits are used to modify the cost-function, so that no SWAP gates are required. In this case QAOA is a continuous-time quantum algorithm.
2. SWAP gates are used. QAOA is then a discrete-time quantum algorithm.

Finally, if the native gate set does not include ZZ, then QAOA is not a continuous-time quantum algorithm.

Divorcing QAOA from native interactions has certain advantages. Auxiliary qubits are not required in the case of finite connectivity. Higher-order cost functions (e.g. ZZZ-rotations) are more straightforward to implement [88]. More complicated driver Hamiltonians that enforce certain constraints can be realised without creating new hardware [89]. Using the driver Hamiltonian to enforce constraints is often referred to as the Quantum Alternating Operator Ansatz [89] (also dubbed QAOA). In conclusion, QAOA is typically a discrete time quantum algorithm.

Other variants of QAOA have been devised, including warm-starting the approach [90]. In general, QAOA suffers from challenges in optimising the free parameters in the QAOA state vector [76, 91, 92]. Modifying the cost function [93] and heuristics [76] have been suggested as a possible way to circumvent this issue [94].

In Chapter 7, QAOA is used as a benchmark and compared to a CTQO

approach. To do this, the runtime of the CTQO approach is compared to the runtime of the QAOA approach. This is done by assuming that QAOA is a CTQO approach, where the runtime of a single instance is taken to be:

$$T_{\text{QAOA}} = \sum_{k=1}^p \gamma_k + \beta_k, \quad (2.25)$$

this excludes the time required to determine suitable choices of γ_k s and β_k s, which would be required in conventional QAOA [14].

2.8 Summary

The CTQO approaches fall into two broad classes. Forward approaches that map from the ground-state of the driver Hamiltonian and cyclic approaches that start in a state diagonal in the computational basis. The approaches are summarised in Table 2.1.

Cyclic approaches are less investigated than forward approaches but present an opportunity to use prior information and to circumvent short coherence times.

Considering the variety of forward approaches, deciding on the best one is a challenge. Brady *et al.* [95, 96, 97] applied optimal control to investigate what the optimal schedule is for a Hamiltonian of the form:

$$H = u(t)H_d + (1 - u(t))H_p, \quad (2.26)$$

where the time T is fixed and $\langle H_p \rangle$ is minimised at the final time. The value of $u(t)$ was bounded between 0 and 1. It was assumed that the minimum of $\langle H_p \rangle$ could not be reached exactly in the given time. Brady *et al.* found that the optimal schedule started with $u(0) = 0$ and finished with $u(T) = 1$. In the middle, the schedule was typically found to be an annealing like schedule in many cases. This optimal schedule was described as a ‘bang-anneal-bang’ schedule. This was later investigated by Venuti *et al.* [98] who provided an extension to open quantum systems.

Algorithm	Abbreviation	Initial state	$A(0)$	$A(T)$	$B(0)$	$B(T)$	$A(t)$	$B(t)$
Adiabatic quantum optimisation	AQO	Ground state of H_d	1	0	0	1	Adiabatic, continuous, monotonically decreasing	Adiabatic, continuous, monotonically increasing
Quantum annealing	QA	Ground state of H_d	1	0	0	1	Continuous, monotonically decreasing	Continuous, monotonically increasing
Continuous-time quantum walk	CTQW	Ground state of H_d	1	1	γ	γ	Constant	Constant
Multi-stage quantum walk	MSQW	Ground state of H_d	1	1	γ_i	γ_f	Constant	Step wise increasing
Reverse quantum annealing	RQA	Eigenstate of H_p	0	0	1	1	Cyclic	Cyclic
Biased quantum annealing*	BQA	Eigenstate of H_p	0	0	1	1	Cyclic	Cyclic
Quantum approximate optimisation algorithm	QAOA	Ground state of H_d	0	1	1	0	$1 - B(t)$	0 or 1

Table 2.1: Summary of CTQO approaches and QAOA. The table summarises the initial state and the schedules $A(t)$ and $B(t)$, including the boundary conditions at the initial time $t = 0$ and the final time $t = T$. The asterisk * denotes the required addition of a third term that biases towards the initial state.

Finding the exact optimal schedule is typically numerically challenging [95].

Work by Gerblich *et al.* [99] demonstrated that MSQWs were closer to the optimal protocol, in terms of operator norm, than QAOA. Lin *et al.* [100] investigated the optimal schedule for the unstructured search problem and found it to be very close to a CTQW. For more general combinatorial optimisation problems, Schulz *et al.* [101] found that given a fixed short runtime on a QPU, CTQWs provide a better time-to-solution than QA, while, given a longer runtime, QA has a better time-to-solution for the problems considered. In summary, there is no definitive answer over which CTQO approach is better.

In this dissertation, we focus on CTQWs in Chapters 4 and 5. Chapter 6 focuses on MSQWs and the generalisation to CTQO approaches more generally, including cyclic approaches. Chapter 7 investigates a novel approach for tackling combinatorial optimisation problems based on Hamiltonians for optimal state transfer.

2.9 A note on numerical work and presentation

This work makes use of numerical experiments to establish the performance of new approaches. The results are often presented as a box-plot [102, 103]. The central line shows the median. The top and bottom line of the central box shows the lower and upper interquartile. Outliers are determined if they are more than 1.5 times more than the interquartile range (IQR) away from the median and are denoted by circles. The two caps at the end of the plot show the maximum and minimum data points, excluding outliers. The box-plot makes no assumption about the underlying distribution and provides a reasonable representation of the distribution for easy comparison. An illustrative box-plot is shown in Fig. 2.3.

The numerical work in this dissertation was completed using the standard numerical simulation tools in the Python package QuTip [104, 105]

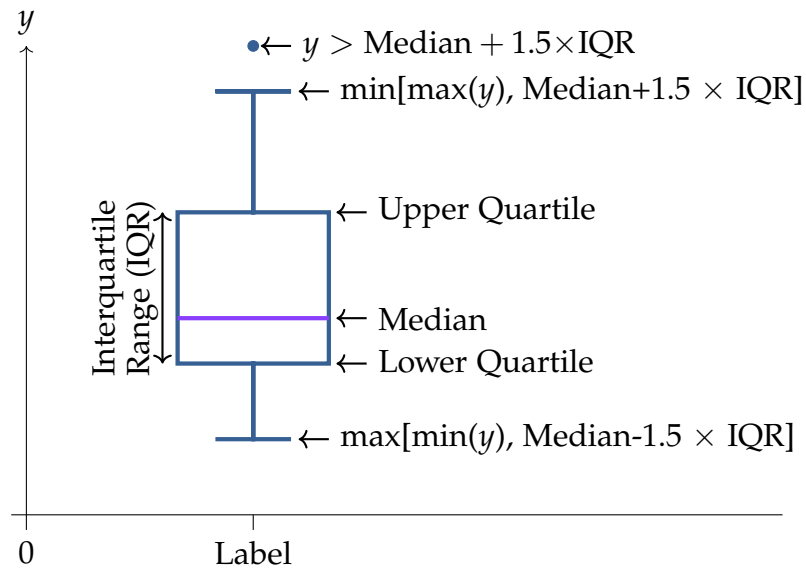


Figure 2.3: An illustration of a box-plot for a data set, named y . The minimum value of y is denoted by $\min(y)$, the maximum value is denoted by $\max(y)$.

and graphs were generated using the Python package NetworkX [106]. Some of the data in Chapter 5 was collected using the Julia package HOQST [107].

Chapter 3

Pure-state statistical physics

3.1 Introduction

The aim of CTQO is to use a many-body quantum system to tackle classical combinatorial optimisation problems. By necessity, exact numerical diagonalisation of the quantum system must be prohibitively difficult. Indeed, for a quantum advantage to exist, it must not be possible for there to be sufficiently good approximate quantum simulators. On the assumption that numerical simulation is infeasible, there needs to be alternative approaches to reasoning about the behaviour of CTQO algorithms. The most common approach has been to appeal to the adiabatic theorem. AQO has its limitations, as discussed in the previous chapter, and therefore there has been a push to non-adiabatic approaches. In this regime, the adiabatic theorem provides little intuition. Alternative approaches to reasoning about CTQO algorithms are therefore required. In this chapter, we demonstrate that pure-state statistical physics can allow for reasoning about CTQO.

Statistical physics has historically proved to be a useful approach to understanding many-body phenomena and making predictions [108]. The framework, at least in the canonical ensemble, involves the identification of a large bath that can exchange energy with the system of interest. Statistical physics is then used to make predictions about that system of interest. Pure-state statistical physics is an attempt to transfer the insights of statistical mechanics to isolated quantum systems. Given a sufficiently large closed quantum system, if the observable of interest

is sufficiently local, can the rest of the system act as its own bath? That is to say, given a closed quantum system, are local observables indistinguishable from a Gibbs state? This was broached in the affirmative for a wide class of systems by Srednicki [23] and Deutsch [24], and has become known as the Eigenstate Thermalisation Hypothesis (ETH). The rest of this chapter outlines some principles that fall under the umbrella of pure-state statistical physics. Some are provably true and some are physically reasonable assumptions based on conceptions of a typical many-body quantum system.

3.2 Conservation of Energy

Energy in a closed-system is conserved: a banal but foundational statement. At a practical level, given a time-independent Hamiltonian H that generates unitary evolution, $\langle H \rangle$ is constant throughout the evolution. The energy provides a conserved quantity that is typically local, that places constraints on the evolution of the system. Many CTQO algorithms, discussed in Chapter 2, have time-dependent Hamiltonians, so energy is not conserved. This means that the systems are not closed, yet it is assumed that the dynamics are unitary. These systems can be described as being isolated.

3.3 Eigenstate Thermalisation Hypothesis

Consider the expectation value of some local observable A in a closed system with an initial density operator ρ and evolved under some local time-independent Hamiltonian [1]. The eigenvectors and associated eigenvalues of the Hamiltonian are denoted by $|E_k\rangle$ and E_k respectively.

It follows that the expectation of A is given by:

$$\langle A(t) \rangle = \sum_{m,n} e^{i(E_m - E_n)t} \langle E_n | \rho | E_m \rangle \langle E_m | A | E_n \rangle \quad (3.1)$$

$$= \sum_{E_n = E_m} \langle E_n | \rho | E_m \rangle \langle E_m | A | E_n \rangle \quad (3.2)$$

$$+ \sum_{E_n \neq E_m} e^{i(E_m - E_n)t} \langle E_n | \rho | E_m \rangle \langle E_m | A | E_n \rangle \quad (3.3)$$

$$= \overline{\langle A \rangle} + \langle \Delta A \rangle, \quad (3.4)$$

where:

$$\overline{\langle A \rangle} = \sum_{E_n = E_m} \langle E_n | \rho | E_m \rangle \langle E_m | A | E_n \rangle \quad (3.5)$$

and

$$\langle \Delta A \rangle = \sum_{E_n \neq E_m} e^{i(E_m - E_n)t} \langle E_n | \rho | E_m \rangle \langle E_m | A | E_n \rangle. \quad (3.6)$$

The time independent sum in Eq. 3.5 corresponds to the infinite-time average of $\langle A(t) \rangle$, i.e.

$$\overline{\langle A \rangle} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle A(t) \rangle dt. \quad (3.7)$$

If $\langle A(t) \rangle$ approaches a steady state, it will approach its infinite-time average. For $\langle A(t) \rangle$ to approach a steady state, $\langle \Delta A \rangle \approx 0$ up to some small fluctuations. Eq. 3.6 provides some intuition for how $\langle A(t) \rangle$ approaches a steady-state – namely, that the difference in energy gaps between pairs of eigenstates will cause dispersion. As a consequence $\langle \Delta A \rangle \approx 0$ up to some negligible fluctuation for large systems, provided the system has little degeneracy. This means that the expectation of A will approach some steady state after some time. This timescale will be problem specific; details on how it might be estimated can be found in [109, 110]. It has been numerically observed and justified that this timescale will not necessarily be exponential in the problem size [109, 111]. This is discussed further for CTQO in Chapter 5. For CTQO the observables that we are most interested in are the problem Hamiltonian H_p and

the driver Hamiltonian H_d . The assumption that dispersion leads to a steady state for local observables is summarised in the two assumptions below.

Assumption 1 (Stationary value) *Under evolution by a time-independent Hamiltonian, the expectation of a local observable A can be replaced with its steady state value $\overline{\langle A \rangle}$ after some time τ_d . This includes the expectation values of H_p and H_d . The time τ_d is the timescale associated with dispersion in the energy basis.*

Assumption 2 (Diagonal in the energy eigenbasis) *Once the system approaches a steady state, we can approximate the state-vector by a density operator which is diagonal in the energy eigenbasis [1].*

The ETH is an attempt to explain how thermalisation occurs in a closed-system. Explicit details can be found in [112]. In this dissertation we are concerned with the *consequences* of the ETH. Primarily, that for a large class of observables, the steady-state value of an observable is equal to averaging over the microcanonical ensemble in the thermodynamic limit. The fluctuations from this value are exponentially suppressed with the degrees of freedom in the system [112]. In certain cases, a single eigenstate is sufficient to carry out the microcanonical averaging [23].

Not all Hamiltonians exhibit ETH, and not all observables thermalise [113]. Typically, the Hamiltonian is assumed to be highly non-degenerate [23, 24, 114] and to provide a locally conserved quantity [115]. The observables that exhibit thermalisation are typically considered to be local or few-body operators [23, 114]. For example, consider the highly non-local operator for a non-trivial system [115]:

$$\mathcal{O}_{jk} = -i (|E_j\rangle \langle E_k| - |E_k\rangle \langle E_j|), \quad (3.8)$$

where $E_j \neq E_k$. This will clearly not approach a steady state.

The ETH tells us that for large systems the steady state, exhibited by

local observables, can be well approximated by the microcanonical ensemble. Though there are known exceptions, this behaviour is thought to be very common [116, 117]. Note that it is possible for a system to approach a steady state that is not a thermal state [115].

For ease of computation, throughout this dissertation we make use of the canonical ensemble, which is equivalent to the microcanonical ensemble in the thermodynamic limit [114, 118]. That is to say, throughout this dissertation we assume the pure quantum state is locally well approximated by a thermal Gibbs state:

$$\rho_{\text{Gibbs}} = \frac{e^{-\beta H}}{\text{Tr } e^{-\beta H}}, \quad (3.9)$$

with temperature $1/\beta$. The temperature is fixed according to the energy of the system:

$$\langle H \rangle = \text{Tr } H \rho_{\text{Gibbs}}. \quad (3.10)$$

This leads to our third assumption:

Assumption 3 (ETH) *The steady state of a local observable is locally indistinguishable from a Gibbs state. The temperature of ρ_{Gibbs} is fixed according to the energy of the system. That is to say, the steady state value $\overline{\langle A \rangle}$ can be well approximated by using the Gibbs state, $\overline{\langle A \rangle} \approx \text{Tr}(A \rho_{\text{Gibbs}})$. The inverse temperature β is fixed by replacing A with the Hamiltonian.*

The ETH, as described above, is likely to be a better approximation for large problem sizes where subextensive corrections and fluctuations can be ignored [119]. A CTQW is a time-independent Hamiltonian and in most settings unlikely to be integrable. We might therefore expect it to exhibit thermalisation for local observables. By exploiting the observation that the system is well approximated by a thermal state, we make predictions on the performance of CTQWs in Chapter 4.

3.4 Spin-flip symmetry and the ETH

This dissertation primarily focuses on Max-cut. A CTQW Hamiltonian, with a transverse-field driver Hamiltonian has a spin-flip symmetry. Let H_{QW} be an example Hamiltonian with spin-flip symmetry, then $[H_{QW}, F] = 0$, where:

$$F = \prod_{i=1}^n X_i. \quad (3.11)$$

Typically, a CTQW starts in the $|+\rangle$ state, so the evolution is restricted to the plus one eigenspace of F for the entire evolution.

Despite this restriction, throughout this dissertation, we use the ‘complete’ Hamiltonian as opposed to the Hamiltonian projected onto the correct symmetry sector to apply the ETH. This is numerically verified in Sec. 4.3. This choice is justified on the following grounds [115]:

1. F is a global operator and corresponds to a global symmetry.
2. We are interested in calculating the expectation values of local interactions only.

It is well established within the ETH literature that local symmetries can prevent thermalisation [115]. In contrast, the spin-flip symmetry is a global symmetry. In general, there are many globally conserved quantities for evolution under a constant Hamiltonian. If H_{QW} belongs to a Hilbert Space \mathcal{H} with dimension $\dim \mathcal{H}$, then there are at least $\dim \mathcal{H}$ globally conserved quantities [115]. To see this, let the projector onto each eigenstate of H_{QW} be denoted by P_k with $k = 1, \dots, \dim \mathcal{H}$, then [115]:

$$[H_{QW}, P_k] = 0 \quad (3.12)$$

$$[P_k, P_j] = 0. \quad (3.13)$$

Hence the $\langle P_k \rangle$ are globally conserved quantities. Despite the large number of conserved quantities, we generally do not expect these global symmetries to have much impact on the dynamics.

Though the above argument suggests we should be careful when considering the role of globally conserved quantities, it fails to consider the specific case of the spin-flip symmetry, which we know restricts dynamics to the positive eigenspace of F for a CTQW. Here we emphasise again that we are only interested in local quantities. A locally measurable consequence of the spin-flip symmetry is that:

$$\begin{aligned}\langle Z_i(t) \rangle &= \langle + | U^\dagger Z_i(t) U | + \rangle \\ &= \langle + | F U^\dagger Z_i(t) U F | + \rangle \\ &= \langle + | U^\dagger F Z_i(t) F U | + \rangle \\ &= - \langle + | U^\dagger Z_i(t) U | + \rangle \\ &= - \langle Z_i(t) \rangle,\end{aligned}$$

where we have used the initial state being an eigenstate of F and $[U, F] = 0$, consequently $\langle Z_i(t) \rangle = 0$.

Since $\langle Z_i(t) \rangle$ is a local observable, this should agree with the result if we were to replace the unitary evolution with a Gibbs distribution:

$$\langle Z_i(t) \rangle_\beta \propto \text{Tr} \left(Z_i e^{-\beta H_{QW}} \right). \quad (3.14)$$

As $[H_{QW}, F] = 0$, this is also zero. Thus, the thermal-state recovers the correct expectation value for any local observable that flips sign under conjugation by F (i.e., $F O_L F = -O_L$).

That is not to say that the spin-flip symmetry has no role in calculating expectation values. Consider the case where the initial state

$$|\psi_i\rangle = \cos \theta/2 |\varphi_+\rangle + e^{-i\varphi} \sin \theta/2 |\varphi_-\rangle, \quad (3.15)$$

is a linear superposition of two states, where $F |\varphi_+\rangle = |\varphi_+\rangle$ and

3. PURE-STATE STATISTICAL PHYSICS

$F|\varphi_-\rangle = -|\varphi_-\rangle$. Then

$$\begin{aligned}\langle\psi_i|U^\dagger O_L U|\psi_i\rangle &= \cos^2\theta/2 \langle\varphi_+|U^\dagger O_L U|\varphi_+\rangle \\ &+ \sin^2\theta/2 \langle\varphi_-|U^\dagger O_L U|\varphi_-\rangle + e^{-i\varphi} \cos\theta/2 \sin\theta/2 \langle\varphi_+|U^\dagger O_L U|\varphi_-\rangle \\ &+ e^{i\varphi} \cos\theta/2 \sin\theta/2 \langle\varphi_+|U^\dagger O_L U|\varphi_-\rangle^*. \quad (3.16)\end{aligned}$$

If O_L also commutes with F then (as is the case with H_p) it follows that:

$$\langle\varphi_+|U^\dagger O_L U|\varphi_-\rangle = -\langle\varphi_+|FU^\dagger O_L UF|\varphi_-\rangle \quad (3.17)$$

$$= -\langle\varphi_+|U^\dagger O_L U|\varphi_-\rangle, \quad (3.18)$$

hence $\langle\varphi_+|U^\dagger O_L U|\varphi_-\rangle = 0$ and

$$\begin{aligned}\langle\psi_i|U^\dagger O_L U|\psi_i\rangle &= \cos^2\theta/2 \langle\varphi_+|U^\dagger O_L U|\varphi_+\rangle \\ &+ \sin^2\theta/2 \langle\varphi_-|U^\dagger O_L U|\varphi_-\rangle. \quad (3.19)\end{aligned}$$

However, we have conjectured (and numerically investigated) that each of the above matrix elements can be represented by replacing each unitary evolution with a thermal state. That is to say,

$$\begin{aligned}\langle\psi_i|U^\dagger O_L U|\psi_i\rangle &\approx \cos^2\theta/2 \frac{\text{Tr}\{O_L e^{-\beta_+ H_{QW}}\}}{\text{Tr}\{e^{-\beta_+ H_{QW}}\}} \\ &+ \sin^2\theta/2 \frac{\text{Tr}\{O_L e^{-\beta_- H_{QW}}\}}{\text{Tr}\{e^{-\beta_- H_{QW}}\}}, \quad (3.20)\end{aligned}$$

therefore the system would need to be assigned two temperatures, one for each symmetry sector.

Throughout this chapter we are interested in $\langle H_p \rangle$, which is invariant under F . Work on spontaneous symmetry breaking and the ETH can be found in [120, 121].

3.5 The second-law of thermodynamics in an isolated-system

So far, the discussion of pure-state statistical physics has been limited to time-independent Hamiltonians. A well-known maxim is that entropy can only increase. This familiar refrain is an expression of the second-law of thermodynamics and places restrictions on the evolution of a system. In this section, we examine similar ideas for pure-state statistical physics.

3.5.1 Diagonal Entropy

Entropy plays a central role in classical statistical physics. In quantum mechanics the von Neumann entropy [82], given by

$$S_{vN}(\rho) = -\text{Tr} \rho \log \rho, \quad (3.21)$$

typically plays this role. Under unitary evolution, the von Neumann entropy is conserved, it is therefore not a useful measure of entropy in an isolated quantum system. A possible candidate is the entanglement entropy. The entanglement entropy is given by tracing out part of a system and taking the von-Neumann entropy of the reduced density operator.

An alternative proposed candidate for the role of entropy in an isolated system has been the diagonal entropy [122]. This is the von Neumann entropy of the diagonal of the density operator in the energy eigenbasis. That is to say, given a Hamiltonian H and associated density operator ρ , the diagonal of ρ is given by

$$\rho_{\text{diag}} = \sum_{\alpha} |E_{\alpha}\rangle \langle E_{\alpha}| \rho |E_{\alpha}\rangle \langle E_{\alpha}|, \quad (3.22)$$

where $|E_{\alpha}\rangle$ are the eigenstates of H . Then the diagonal entropy of ρ is

given by:

$$S_d(\rho) = -\text{Tr} \rho_{\text{diag}} \log \rho_{\text{diag}}. \quad (3.23)$$

The diagonal entropy is conserved in a closed system. This is not necessarily true for isolated systems. From Jensen's inequality, it is known that [122]:

$$S_d(t) \geq S_{vN}, \quad (3.24)$$

where $t \geq 0$ denotes the time of evolution. If $S_d(0) = S_{vN}$, meaning that the initial state is diagonal in the energy eigenbasis, then Eq. 3.24 implies [122]:

$$S_d(t) \geq S_d(0). \quad (3.25)$$

Hence, if the initial state is diagonal in the energy eigenbasis, then the diagonal entropy can only increase compared to its initial value. This does not imply $S_d(t)$ is monotonically increasing in time. Further to this, if the system has approached a steady state (Assumption 2) and is well approximated by a diagonal state, then it might reasonably be expected that varying the Hamiltonian will cause the diagonal entropy to increase. Although the diagonal entropy does not capture the strict monotonicity desired, it does reasonably approximate this behaviour. Further to this, in AQO/QA excitations to higher energy states are primarily of interest and captured by the diagonal entropy.

3.5.2 Planck's Principle

Calculating the diagonal entropy requires knowledge of ρ_{diag} . Typically, we are interested in observables such as $\langle H_p \rangle$, while diagonal entropy does not correspond to an observable. For this reason, a further assumption is introduced that relates schedules to an observable — namely the energy.

Assumption 4 (Planck's Principle) *For any relevant cyclic process in an isolated system, no work can be extracted. That is, for any cyclic process represented by a unitary U , with initial Hamiltonian H and initial state ρ , the*

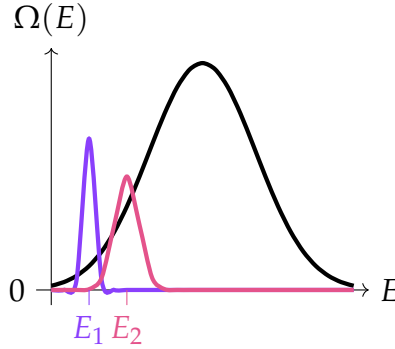


Figure 3.1: The solid black line sketches the typical density-of-states $\Omega(E)$, for a large non-integrable system [1]. The solid blue line shows a possible energy distribution for the initial state. The red line shows a possible resulting energy distribution as the result of a cyclic process. The energy distribution has been broadened and has moved towards the middle of the spectrum.

following holds true:

$$W = E_{Initial} - E_{Final} = \text{Tr} \left[H \left(\rho - U\rho U^\dagger \right) \right] \leq 0. \quad (3.26)$$

This assumption has significant consequences for CTQO. To justify it, we show its consistency with Assumptions 3 and 2. We also point the reader to [26] where Goldstein *et al.* showed that under certain assumptions, in the isolated system setting, the extractable work is exponentially more likely to decrease than not. The assumptions include that knowledge of the initial state (i.e. before the cyclic process) is lost by time-averaging under evolution by a Hamiltonian — which is the case in an MSQW. Assumption 4 is sometimes referred to as Kelvin’s statement of the second law of thermodynamics [1], or Planck’s principle [26, 27]. See also [1, 27, 123, 124, 125] for further attempts to mathematically motivate the second law of thermodynamics in isolated systems from the rules of quantum mechanics.

It is known that any Gibbs state, with a positive temperature, satisfies Eq. 3.26 [126]. A very simple derivation is as follows: let the initial state ρ_0 be a Gibbs state at inverse temperature $\beta > 0$ with Hamiltonian H . The state after unitary evolution is denoted by $\rho(t)$. The Gibbs state

minimises the free energy:

$$\mathrm{Tr}\{H\rho_0\} - \frac{1}{\beta} S_{vN}(\rho_0) \leq \mathrm{Tr}\{H\rho(t)\} - \frac{1}{\beta} S_{vN}(\rho(t)) \quad (3.27)$$

where $S_{vN}(\cdot)$ is the von Neumann entropy. Rearranging the above gives:

$$S(\rho(t)) - S(\rho_0) \leq \beta (\mathrm{Tr}\{H\rho(t)\} - \mathrm{Tr}\{H\rho_0\}). \quad (3.28)$$

Since the von Neumann entropy is conserved under unitary evolution,

$$0 \leq \beta (\mathrm{Tr}\{H\rho(t)\} - \mathrm{Tr}\{H\rho_0\}). \quad (3.29)$$

Provided that the temperature is positive,

$$\mathrm{Tr}\{H\rho_0\} \leq \mathrm{Tr}\{H\rho(t)\}. \quad (3.30)$$

Hence for any unitary, $\langle H \rangle$ can only increase for a Gibbs state. This means that, for a cyclic process, the energy can only increase. This means that Planck's principle is provably true for Gibbs states. A Gibbs state is an example of a passive state. A passive state is a state that is diagonal in the energy eigenbasis of a Hamiltonian, with the populations ordered such that they do not increase with energy [126]. Examples include ground states and sampling energy eigenstates from a uniform distribution with only a high-energy cut-off. All passive states provably satisfy Planck's principle [127].

If we assume that the ETH (i.e. Assumption 3) holds for the initial state, then it remains to determine if $U^\dagger H U$ is a local observable [1, 119]. If $U^\dagger H U$, where U corresponds to a cyclic process, is a local observable, then Assumption 4 follows as a consequence of Assumption 3.

For further evidence of Planck's principle, away from the ETH, we might look to the diagonal entropy. Coupled with a reasonable unimodal model of the density-of-states (especially for non-integrable systems) [1], as sketched in Fig. 3.1, this implies Assumption 4. We expect this

to hold for initial Hamiltonians like H_p too (see for example [101]) for a wide range of problems. Essentially, under a cyclic process, the system is expected to heat up as there are more states towards the middle of the spectrum than at the edges. Under continued periodic drive, this is sometimes referred to as Floquet heating [1, 128]. In order to apply Assumption 4, it is required that the system is more likely to move up in energy than down in energy. This is unlikely to be the case given an initial state that is very high in energy (corresponding to a negative temperature Gibbs state). Typically, the system is expected to move towards the infinite temperature state as a result of a cyclic process. This places a restriction on the initial state, $|\psi_i\rangle$, for Assumption 4 to hold:

$$\langle \psi_i | H | \psi_i \rangle < \frac{1}{\mathcal{D}} \text{Tr } H, \quad (3.31)$$

where \mathcal{D} is the dimension of the Hilbert space. Perhaps a more accurate statement of Assumption 4 is that a cyclic process will move the energy of the system towards its infinite temperature value [1, 128]. For the purpose of what is to follow, Assumption 4 as stated originally is sufficient. Although there are specific cases which violate each assumption we expect them to hold true for a wide variety of problems, drivers and encodings, especially as the problem size is increased.

Assumptions 2, 3 and 4 all have the effect of imposing an arrow of time on the evolution. This comes from discarding the coherences in the energy eigenbasis. In a closed system, this is a result of dispersion in the energy eigenstates. For an open system, this could come from being very weakly coupled to a bath. Notably, AQO makes no use of coherences in the energy eigenbasis. We will show that the performance of MSQWs does not rely on coherence between energy eigenstates either — despite being non-adiabatic. It remains an open question, how (or if) coherences in the energy eigenbasis in CTQO can be (or should be) reliably exploited.

3.6 Conclusion

Pure-state statistical physics is largely intuitive. Planck's principle is based on the assumption there are more states towards the middle of the energy spectrum and less at the edge, therefore a cyclic process is more likely to increase in energy. Given the energy of a system in a steady-state, a sensible estimator of the state is the one that maximises the entropy (i.e. the Gibbs state). Pure-state statistical physics is not without its limitations. In many regimes, the results are hypotheses as opposed to provable statements. Planck's principle does not always hold, nor is the ETH always true.

Given a closed system, there are numerous observables that are conserved, for example all powers of the Hamiltonian. Therefore, a more sensible estimator of the state might be a generalised Gibbs ensemble, with a Lagrange multiplier for each power of H . Each power of H provides a less local conserved quantity (i.e. operators with larger support). By taking a Gibbs state, the assertion is that the most local conserved quantity is the most important and the rest can be neglected. In this dissertation, it is shown that in many cases the Gibbs state is sufficient to make predictions. The Gibbs state has the additional benefit of being a well-studied state. There are current attempts to improve on the ETH [129, 130, 131, 132, 133].

This dissertation uses the framework laid out in this chapter to gain intuition about the CTQO algorithms discussed in Chapter 2. This is often explicitly numerically verified for small problem instances where exact numerical diagonalisation is achievable. Given that the arguments rely on the intuition contained in this chapter, it is expected that the conclusions drawn will hold outside these explicitly verifiable problem instances.

Chapter 4

Continuous-time quantum walks for Max-cut are hot

4.1 Introduction

Continuous-time quantum walks (CTQWs) are a relatively new approach to tackling combinatorial optimisation problems. This chapter discusses how to reason about CTQWs where exact numerical diagonalisation is prohibitively difficult. To discuss CTQWs, this chapter focuses on Max-cut due to its simple Ising representation. Max-cut, along with the graph types considered in this chapter, were introduced in Sec. 2.3. The driver Hamiltonian is taken to be the standard transverse-field driver. In a CTQW, the initial state is evolved under a time-independent Hamiltonian, with some free parameter γ . Sec. 4.2 investigates optimising γ numerically for small problem instances. As discussed in Chapter 3, it is expected that CTQWs will thermalise. In Sec. 4.3 it is numerically verified that this is the case. The rest of this chapter focuses on the long-time behaviour of CTQWs. Discussion of the dynamics is postponed until Chapter 5. Sec. 4.4 provides two analytically tractable models for the density-of-states (DOS). The models are then used to make predictions about the steady-state behaviour of CTQWs (Sec. 4.5), including the optimal choice of γ in Sec. 4.6. In Sec. 4.7 some evidence of the performance of optimised CTQWs is provided.

4.2 Optimising the free parameter

The CTQW Hamiltonian

$$H_{QW} = H_d + \gamma H_p \quad (4.1)$$

contains a free parameter γ . Heuristically, this controls the amount of dynamics present in the evolution. In this section, the free parameter γ is optimised numerically to give the best time-averaged value of $\langle H_p(t) \rangle$. More explicitly, $\langle H_p(t) \rangle$ is given by

$$\langle H_p(t) \rangle = \langle + | e^{iH_{QW}t} H_p e^{-iH_{QW}t} | + \rangle \quad (4.2)$$

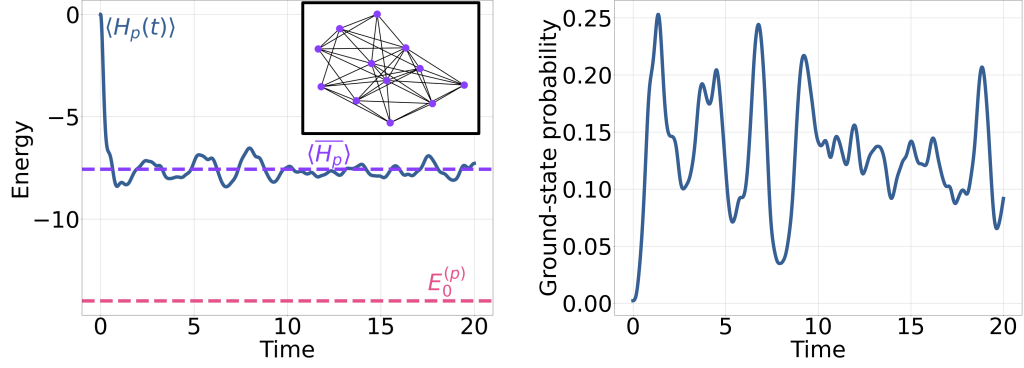
and the time-averaged value is given by:

$$\overline{\langle H_p \rangle} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle H_p(t) \rangle dt. \quad (4.3)$$

In the rest of this section, by brute force optimisation, the optimal choice of γ is found for different graph choices. Comparisons to a reasonable heuristic choice of γ are also made. To begin with, the specific example of a 12-qubit binomial graph (shown in the inset of Fig. 4.1a) is focussed on, before looking at statistics gathered from multiple graph instances.

We start with a specific choice of $\gamma = 1$ for the graph instance shown in the inset of Fig. 4.1a. The blue line in Fig. 4.1a shows the average value of H_p for the CTQW on this graph with $\gamma = 1$. The dashed pink line shows the ground state energy of H_p (i.e. $E_0^{(p)}$). For the majority of the evolution shown $\langle H_p(t) \rangle$ is fluctuating around the steady-state value, $\overline{\langle H_p \rangle}$ (the dashed purple line). The ground state probability is shown in Fig. 4.1b. The ground state probability shows significant oscillations, although these oscillations may have benefits in solving the optimisation problem, the steady-state behaviour of $\langle H_p \rangle$ makes it easier to optimise over to find the optimal γ .

Fig. 4.2 shows how $\overline{\langle H_p \rangle}$ varies with γ for the same problem instance



(a) A time domain plot of $\langle H_p(t) \rangle$ (the blue line). The dashed pink line shows the minimum energy of H_p . The dashed purple line is the time-averaged value.

(b) A time domain plot of the ground-state probability

Figure 4.1: Plots of possible metrics for assessment of CTQWs for the randomly generated 12-qubit instance, shown in the inset of Fig. 4.1a

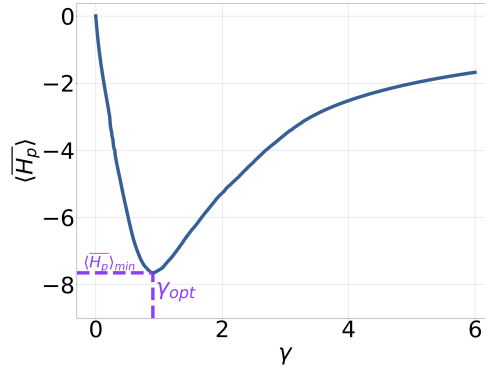


Figure 4.2: The time-averaged value $\overline{\langle H_p \rangle}$ as γ is changed for the 12 qubit graph shown in the inset of Fig. 4.1a. The dashed purple line shows the location of γ_{opt} and $\overline{\langle H_p \rangle}_{min}$.

as Fig. 4.1a. The optimal γ (i.e. γ_{opt}) occurs at the minimum value of $\overline{\langle H_p \rangle}$ (i.e., $\overline{\langle H_p \rangle}_{min}$). For the problem instance considered in Fig. 4.2, $\gamma_{opt} \approx 0.90$ and $\overline{\langle H_p \rangle}_{min} \approx -7.65$.

In previous works, the free parameter in a CTQW has been heuristically chosen in an attempt to maximise dynamics. One approach to do this has been to match the energy scales of the driver and problem Hamiltonian [22, 134]. For a Max-cut problem, this could be interpreted as

$\text{Tr } \gamma^2 H_p^2 = \text{Tr } H_d^2$ such that:

$$\gamma_{\text{heur}}^2 \kappa_2 = n, \quad (4.4)$$

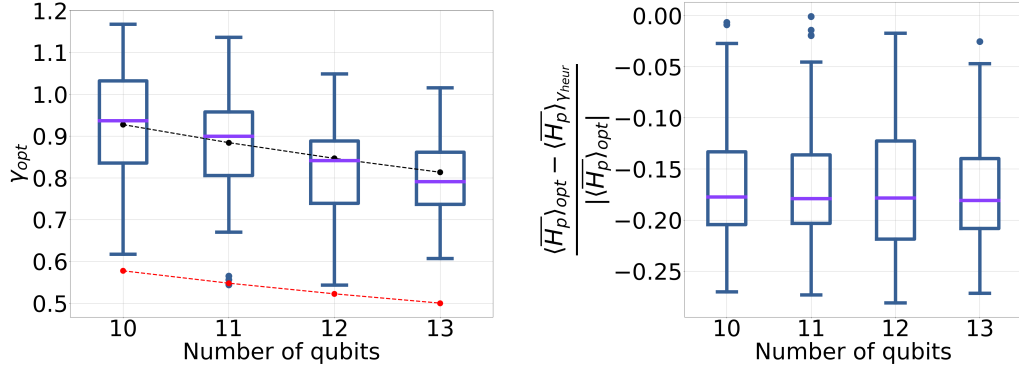
where $\kappa_2 = |E|$ is the number of edges in the graph and n is the number of vertices in the graph and equivalently the number of qubits involved in the CTQW. Focusing first on the performance of the binomial graph shown in Fig. 4.2, evaluating Eq. 4.4 gives $\gamma_{\text{heur}} = 0.5$, corresponding to $\overline{\langle H_p \rangle}_{\min} \approx -5.88$. This is a significant worsening of the performance of the CTQW.

Fig. 4.3a shows the optimal γ for one hundred randomly generated problem instances for binomial graphs with problem sizes ranging between ten and thirteen qubits. The optimal γ is seen to decrease with the problem size. The range of optimal values of γ is on the order of 0.1. The dashed red line represents the heuristic in Eq. 4.4 assuming the number of edges is $n(n-1)/3$ as the probability of selecting an edge is $2/3$ for the graphs generated. Clearly, Eq. 4.4 is significantly underestimating the optimal γ .

It is expected that the optimal γ will balance H_p and H_d , hence why it is reasonable to assume $\gamma \propto n^{-1/2}$. This is the same functional dependence on n as Eq. 4.4. The dashed black line in Fig. 4.3a shows $\gamma \propto n^{-1/2}$ fitted to the available data. The curve is not inconsistent with the data.

Fig. 4.3b shows how $\overline{\langle H_p \rangle}$ changes between γ_{opt} and γ_{heur} , normalised by the value of $\overline{\langle H_p \rangle}$ at the optimal choice of γ for the same problem instances as Fig. 4.3a. The difference in performance is around 15-20%. From the numerical data collected, this difference in performance appears largely independent of problem size. There is clearly scope for improvement on the heuristic choice of γ .

Turning to three-regular graphs, Fig. 4.4a shows γ_{opt} for different problem sizes. For regular graphs, the number of edges for a given problem size is fixed and scales with n . Hence, γ_{heur} evaluates to be the same for all instances of a d -regular graph. This is the red dashed line in the



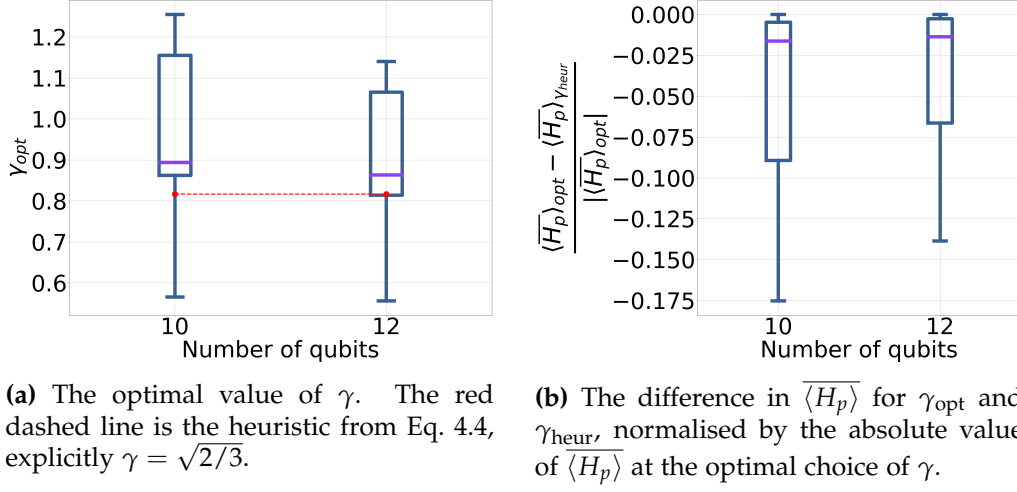
(a) The optimal value of γ . The red dashed line reflects the heuristic from Eq. 4.4, explicitly $\gamma = \sqrt{3/(n-1)}$. The dashed black line corresponds to fitting the curve $\gamma_{\text{opt}} = an^{-1/2}$ to the medians of the data, yielding $a \approx 2.93$.

(b) The difference in $\langle H_p \rangle$ for γ_{opt} and γ_{heur} , normalised by the absolute value of $\langle H_p \rangle$ at the optimal choice of γ .

Figure 4.3: A comparison of the optimised value of γ , i.e., γ_{opt} and the heuristic choice of γ shown in Eq. 4.4 for 100 binomial graphs at each problem size. The number of qubits n is equal to the number of vertices in the Max-cut graph.

figure. For three-regular graphs γ_{heur} also appears to be generally underestimating γ_{opt} but less so than in the binomial case. This is reflected in the change in performance, as shown in Fig. 4.4b.

This section has focused on finding the optimal γ . The closer γ is to the optimal value, the better the approximation ratio is. To find the optimal γ brute-force search was utilized. This could be realised for larger problem sizes by sampling from a device capable of implementing Eq. 4.1, however this approach occurs a sampling overhead. In this section we have focused on exploring a heuristic choice based on the mantra of maximising dynamics for selecting γ . The heuristic choice of γ has elucidated how the optimal choice of γ might scale with n . This is particularly pertinent to the discussion in Sec. 4.4. Sec. 4.6 shows how the heuristic choice of γ (i.e. Eq. 4.4) can be recovered assuming thermalisation and how it might be improved upon. An alternative approach to estimating the optimal γ could have been to extrapolate from small problem instances, a similar approach has been explored in the context of QAOA [135]. The aim of the rest of this chapter is to start to



(a) The optimal value of γ . The red dashed line is the heuristic from Eq. 4.4, explicitly $\gamma = \sqrt{2/3}$.

(b) The difference in $\langle H_p \rangle$ for γ_{opt} and γ_{heur} , normalised by the absolute value of $\langle H_p \rangle$ at the optimal choice of γ .

Figure 4.4: A comparison of the optimised value of γ , i.e., γ_{opt} and the heuristic choice of γ shown in Eq. 4.4 for three-regular graphs. In ascending problem size, there are: 16, 45 problem instances.

understand how to describe CTQWs, with the ultimate aim of setting γ .

4.3 Continuous-time quantum walks are well modelled by thermal states

As described in Chapter 3, a closed quantum system under evolution by a constant Hamiltonian can result in a state that is locally indistinguishable from a thermal state. In terms of CTQWs, of primary interest is $\langle H_p \rangle$, a sum over local observables. Thus, if a CTQW close to the optimal γ exhibits thermalisation, then $\langle H_p \rangle$ should correspond to that of a thermal state. In this section, numerical evidence that CTQWs do exhibit thermalisation even for small problem sizes is provided.

Since the Hamiltonian is constant during the evolution of a CTQW, energy is conserved. Hence, the following must hold true for the thermal state [115]:

$$\text{Tr} [H_{QW} \rho_\beta (H_{QW})] = -n, \quad (4.5)$$

where $\rho_\beta(H_{QW})$ denotes a thermal state with Hamiltonian H_{QW} and inverse temperature β . This equation fixes β , therefore there are no

4.3. Continuous-time quantum walks are well modelled by thermal states

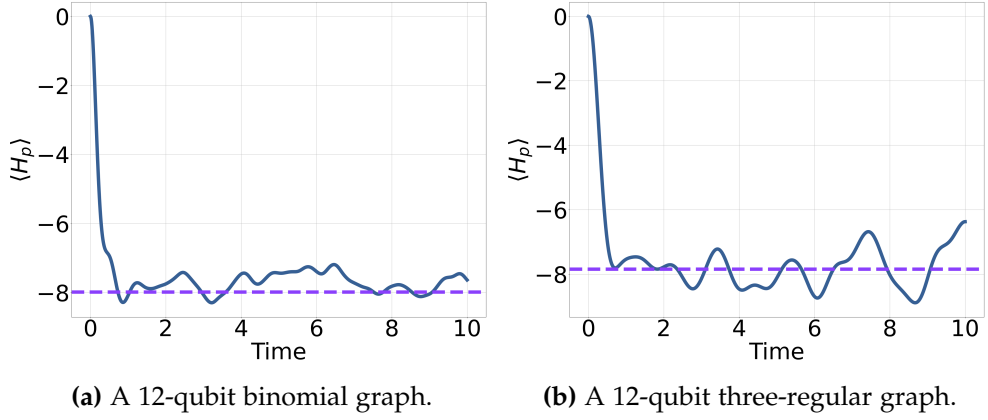


Figure 4.5: The performance of a CTQW with γ optimised to give the best value of $\overline{\langle H_p \rangle}$. The dashed purple line shows the thermal state prediction. The temperature is fixed using Eq. 4.5.

free parameters to fit. Knowing the temperature associated with the CTQW, provides us with a second route to find the same value of $\overline{\langle H_p \rangle}$ by preparing a thermal state, thus side-stepping the need for complete unitary dynamics.

Fig. 4.5 shows the Schrödinger evolution of $\langle H_p \rangle$ for a binomial graph and a three-regular graph and the corresponding prediction for $\langle H_p \rangle$ from the thermal state. For both problem instances, it appears that $\langle H_p \rangle$ is fluctuating around a steady state value. Importantly, despite being far from the thermodynamic limit at only 12 qubits, the thermal state prediction is capturing the steady-state behaviour well.

Fig. 4.6 shows the solution to Eq. 4.5 for the optimal choice of γ for multiple instances of binomial and three-regular graphs. For both cases the inverse temperatures are quite small, with $\beta < 1$ for almost all instances. This means that CTQWs correspond to Gibbs states with high temperatures.

As mentioned in Chapter 3, it is expected that the ETH holds true in the thermodynamic limit. The system sizes simulated here are of the order of ten qubits, hence it is expected that there will be finite-size effects, meaning there will be some error between the Schrödinger equation

4. CONTINUOUS-TIME QUANTUM WALKS FOR MAX-CUT ARE HOT

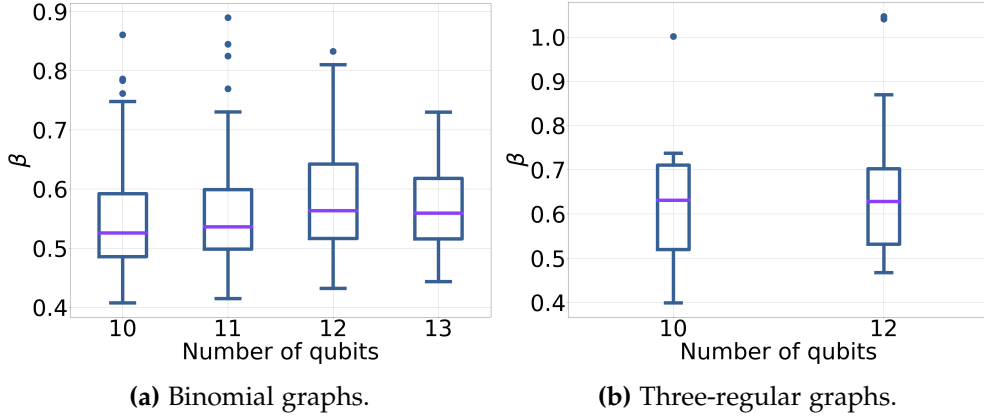


Figure 4.6: The inverse temperature associated with CTQWs (i.e. the solution to Eq. 4.5). For each problem instance γ is equal to γ_{opt} , as shown in Fig. 4.3a and Fig. 4.4a.

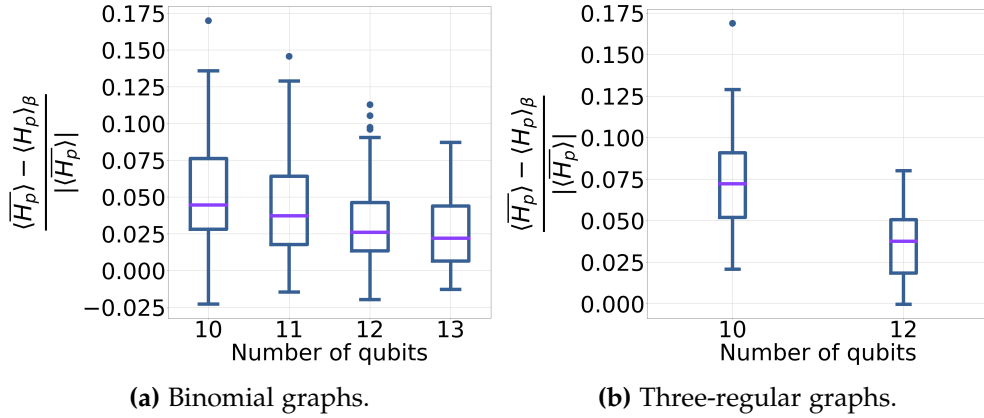


Figure 4.7: The difference between the true value of $\overline{\langle H_p \rangle}$ and the thermal prediction (for the values of β shown in Fig. 4.6) normalised by the magnitude of the true value of $\overline{\langle H_p \rangle}$ for each problem.

and the thermal prediction. This is captured in Fig. 4.7. Generally, the thermal prediction, $\langle H_p \rangle_\beta$ is overestimating the performance of the CTQW. As n is increased, it is expected that this error will decrease. However, even for these very small systems the error is relatively small, with no significant outliers in this data set.

4.4 Modelling the density of states

In the previous section it was numerically demonstrated that, even for relatively small systems, $\overline{\langle H_p \rangle}$ is well predicted by a thermal state. Numerically solving Eq. 4.5 is however difficult, requiring finding the exponential of a matrix that increases exponentially with the problem size. The aim is to find a method for estimating the temperature associated with a CTQW. To do this, possible models for the DOS of H_{QW} are investigated in this section.

To better understand the DOS, it is possible to calculate moments of the distribution. These can be used as fitting parameters for a model of the DOS. To summarise, the goal is to calculate the moments of the distribution produced by the eigenenergies of the CTQW Hamiltonian for the graph $G = (V, E)$:

$$H_{QW} = H_d + \gamma H_p, \quad (4.6)$$

where:

$$H_d = - \sum_{i=1}^n X_i \quad (4.7)$$

$$H_p = \sum_{(i,j) \in E} Z_i Z_j, \quad (4.8)$$

where $n = |V|$ is the number of nodes in the graph. Note that the Hamiltonian is constructed of Pauli matrices, which are traceless [82]. Denoting the eigenenergies of H_{QW} as E_k , then the mean of the distribution of eigenenergies is given by:

$$\mu = \frac{1}{2^n} \sum_k E_k = \frac{1}{2^n} \text{Tr } H_{QW}. \quad (4.9)$$

Since the Hamiltonian is traceless, this evaluates to zero. Noting that only terms that are equal to the identity will contribute to the trace,

repeating this approach for the variance gives:

$$\begin{aligned}
\sigma^2 &= \frac{1}{2^n} \sum_k E_k^2 \\
&= \frac{1}{2^n} \text{Tr } H_{QW}^2 \\
&= \frac{1}{2^n} \text{Tr} \left(H_d^2 + 2\gamma H_d H_p + \gamma^2 H_p^2 \right) \\
&= n + 0 + \gamma^2 \kappa_2,
\end{aligned}$$

where κ_2 is the number of edges in the graph. The same approach can be used to find the third moment:

$$\begin{aligned}
\frac{1}{2^n} \sum_k E_k^3 &= \frac{1}{2^n} \text{Tr } H_{QW}^3 \\
&= \frac{1}{2^n} \text{Tr} \left(\gamma^3 H_p^3 \right) \\
&= 6\gamma^3 \kappa_3,
\end{aligned}$$

where κ_3 is the number of triangles in the graph [136]. The fourth moment also follows from the above logic:

$$\begin{aligned}
\frac{1}{2^n} \sum_k E_k^4 &= \frac{1}{2^n} \text{Tr } H_{QW}^4 \\
&= \frac{1}{2^n} \text{Tr} \left(H_d^4 + 4\gamma^2 H_d^2 H_p^2 + 2\gamma^2 H_d H_p H_d H_p + \gamma^4 H_p^4 \right) \\
&= \left(n^2 + 2(n^2 - n) \right) + 4\gamma^2 n \kappa_2 + 2\gamma^2 (n - 4) \kappa_2 \\
&\quad + \gamma^4 (\kappa_2 + 3\kappa_2(\kappa_2 - 1) + 24\kappa_4),
\end{aligned}$$

where κ_4 is the number of squares in the graph. From here it is straightforward to calculate the skewness and kurtosis. In theory, this process of calculating moments could be continued to higher orders, incorporating loops of greater lengths. However, as shown above, even by the fourth order this becomes cumbersome.

To make further progress, it is assumed that the DOS is well modelled

by a continuous distribution. This assumption is justified because the eigenstates of H_{QW} are likely to become exponentially close as the system size grows. This follows from the largest eigenvalue of H_{QW} being polynomial in n , and there being exponentially many states. Assuming no (or little) degeneracy, the difference between energy levels must shrink exponentially with the problem size. We also assume that the DOS is well modelled by a uni-modal distribution. Intuitively, for large systems, it is expected that the energy of a randomly chosen eigenstate will be close to the average energy. The average energy, μ , of H_{QW} is 0. The variance, σ^2 is equal to $n + \gamma^2 \kappa_2$. Therefore, as n goes to infinity, $\sigma/2^n$ tends to zero. So we expect a tightly peaked distribution for large n . The assumption of uni-modality will break down if γ is too large or too small. However, as seen in Sec. 4.2 we expect useful values of γ to correspond to somewhere between these two limits.

The DOS for thermalising systems has previously been modelled with a Gaussian distribution [1, 137, 138, 139]. Since the CTQW exhibits thermalisation, we adopt this approach. Fig. 4.8 shows a Gaussian fit to the DOS for a 12-qubit three-regular graph (Fig. 4.8a) and a binomial graph (Fig. 4.8b) with γ optimised for each problem. Visually, this appears to be an acceptable approximation for the regular graph, perhaps less so for the binomial graph. Given the simplicity of this approximation, it is utilised throughout the chapter to make simple analytic predictions.

To verify the Gaussianity of the DOS, two typical tests of normality, the skewness, s , of the distribution and the excess kurtosis k , are evaluated. The skewness is the ratio of the third moment of the distribution to the cube of the standard deviation. The kurtosis is the ratio of the fourth moment of the distribution to the variance squared. The excess kurtosis is the kurtosis minus the expected kurtosis of a Gaussian distribution, which is equal to 3.

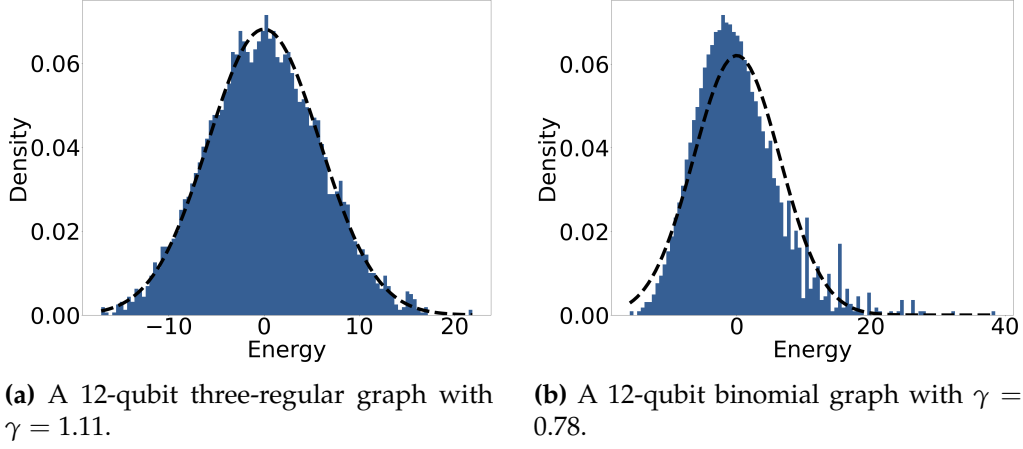


Figure 4.8: Histogram for the DOS for two graphs with optimised γ . The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution.

Starting with the skewness:

$$s = \frac{6\gamma^3\kappa_3}{(n + \gamma^2\kappa_2)^{3/2}}. \quad (4.10)$$

Note that triangle-free graphs have no skewness and that for positive γ , the skewness is always non-negative.

For a d -regular graph, $\kappa_2 = dn/2$ and κ_3 is bounded by $\mathcal{O}(d^2n)$. So the skewness, following the heuristic in Eq. 4.4 (provided γ is held constant) scales approximately as $n^{-1/2}$. Therefore, it will tend to zero as the problem size is increased.

Conversely, for binomial graphs, provided γ scales proportional to $1/\sqrt{n}$, then the skewness will not scale with n . Hence, the Gaussian approximation will not hold as well for a binomial graph.

Examining now the excess kurtosis gives:

$$k = -2 \frac{n + 4\gamma^2\kappa_2 + \gamma^4(\kappa_2 - 12\kappa_4)}{(n + \gamma^2\kappa_2)^2}. \quad (4.11)$$

Under the same assumptions above for the scaling of the skewness, the

excess kurtosis will tend to zero as n tends to infinity for regular graphs. For binomial graphs, the excess kurtosis is unlikely to vanish.

For large regular graphs, the Gaussian approximation looks to hold well. The same cannot be said to be true for binomial graphs. This is a consequence of regular graphs looking locally tree-like in the infinite-size limit [140]. Hence, it may be necessary to consider other models for the DOS for binomial graphs and for regular graphs away from the infinite-size limit.

Here, we propose using an exponentially modified Gaussian (EMG) distribution [141, 142] to model the DOS for a CTQW. The EMG is a skewed unimodal distribution. The skewness can only be positive, as is the case with the DOS for a CTQW. In the correct limit, it can recover a Gaussian. It has the following probability density function:

$$p(x; m, v, \lambda) dx = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2m + \lambda v^2 - 2x)} \operatorname{erfc} \left(\frac{m + \lambda v^2 - x}{\sqrt{2}v} \right) dx, \quad (4.12)$$

where erfc is the complementary error function. The fitting parameters m , v and λ can be ascertained from n , σ and s . The EMG distribution is a convolution of a Gaussian distribution with an exponential distribution [143]. Denoting the exponential distribution as:

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \lambda e^{-\lambda x} & \text{for } x \geq 0 \end{cases} \quad (4.13)$$

and the normal distribution as,

$$g(x) = \frac{1}{\sqrt{2\pi}v} e^{-\frac{(x-m)^2}{2v^2}}, \quad (4.14)$$

then the exponentially modified Gaussian distribution is given by:

$$\begin{aligned}
h(x) &= f(x) * g(x) \\
&= \int_{-\infty}^{\infty} dy f(y)g(x-y) \\
&= \frac{\lambda}{2} e^{\frac{\lambda}{2}(2m+\lambda v^2-2x)} \operatorname{erfc}\left(\frac{m+\lambda v^2-x}{\sqrt{2}v}\right). \tag{4.15}
\end{aligned}$$

Through application of the convolution theorem [144] it is straightforward to write down the moment generating function (and hence partition function) from the moment generating functions of the exponential and Gaussian distributions. The resulting partition function in terms of the fitting parameters m , v^2 and λ is:

$$\mathcal{Z}(\beta) = \left(1 + \frac{\beta}{\lambda}\right)^{-1} e^{-m\beta + \frac{1}{2}v^2\beta^2} \tag{4.16}$$

To find the fitting parameters we fit the moments of $h(x)$ to the moments associated with H_{QW} . Utilising the moment generating function to find the mean (μ), variance (σ^2) and skewness (s) of the EMG gives:

$$\mu = m + \frac{1}{\lambda} \tag{4.17}$$

$$\sigma^2 = v^2 + \frac{1}{\lambda^2} \tag{4.18}$$

$$s = \frac{2}{v^3\lambda^3} \left(1 + \frac{1}{v^2\lambda^2}\right)^{-3/2}. \tag{4.19}$$

The above equations can be inverted to find the fitting parameters in terms of the mean, variance, and skewness associated with H_{QW} . These properties have already been calculated.

The Gaussian approximation only includes information about the problem size, and therefore will make the same prediction for numerous graphs. For instance, all regular graphs with the same degree have the same DOS under the Gaussian approximation. By incorporating the skewness into the model, more information about the graph structure

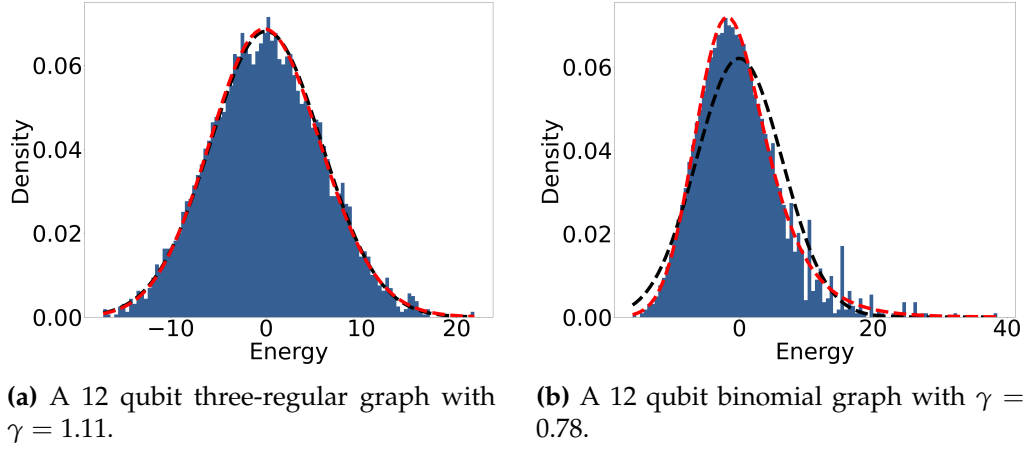


Figure 4.9: Histogram for the DOS for two graphs with γ optimised. The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution. The dashed red line shows the EMG distribution.

is being incorporated, namely the number of triangles in the problem. Visually, as shown in Fig. 4.9, it is clear that the EMG distribution (the dashed red line) provides a better fit than the Gaussian distribution for the binomial graph. The EMG distribution also still models the DOS for the regular graph well.

During the rest of this chapter, the DOS is used to make analytic predictions about the behaviour of a CTQW in a closed-system setting.

4.5 Predictions from the density of states

Having provided a model for the DOS for H_{QW} , in this section two predictions in relation to CTQWs are made, namely the temperature and the diagonal entropy.

Estimating the temperature

Finding the temperature of the CTQW requires finding the solution to Eq. 4.5. To find approximate solutions, we use the DOS models to

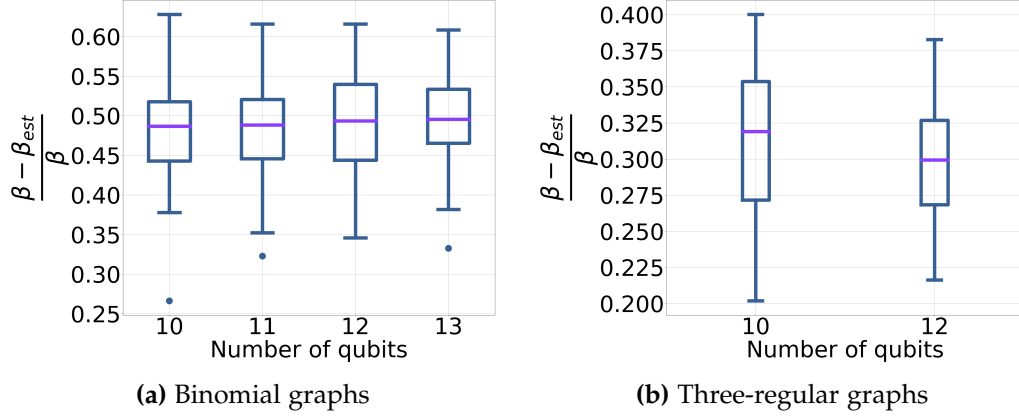


Figure 4.10: The error in the predicted inverse temperature assuming a Gaussian DOS (i.e. Eq. 4.21). For each problem instance γ is equal to γ_{opt} , shown in Fig. 4.3a and Fig. 4.4a.

evaluate the partition function, \mathcal{Z} . Eq. 4.5 then becomes [108]:

$$-n = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}. \quad (4.20)$$

Assuming a Gaussian DOS, then the estimated inverse temperature is given by:

$$\beta_{\text{est}} = \frac{n}{n + \gamma^2 \kappa_2}. \quad (4.21)$$

This is perhaps what one would estimate as the inverse temperature on dimensional grounds alone.

Taking the EMG to model the density of states gives:

$$\beta_{EM} = \frac{n\Delta - \sigma^2 + \sqrt{(n\Delta + \sigma^2)^2 - 4n\Delta^3}}{2\Delta(\sigma^2 - \Delta^2)}, \quad (4.22)$$

where $\sigma^2 = n + \gamma^2 \kappa_2$ and $\Delta = \gamma(3\kappa_3)^{1/3}$.

Fig. 4.10 shows the error in temperature between β_{est} and β (the numerical solution to Eq. 4.5). The approximate solutions are consistently underestimating the inverse temperature. The error for binomial graphs is substantial, with typical errors being between 45% and 50% (Fig.

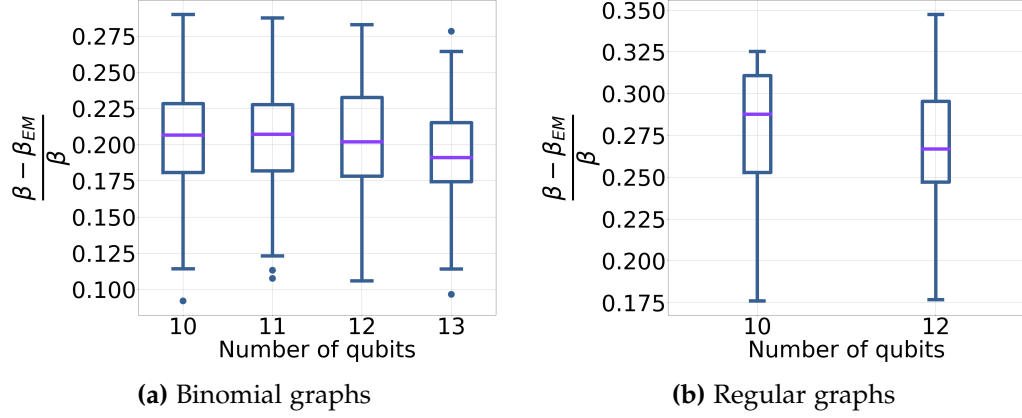


Figure 4.11: The error in the predicted inverse temperature assuming an EMG DOS (i.e. Eq. 4.22). For each problem instance γ is equal to γ_{opt} , shown in Fig. 4.3a and Fig. 4.4a.

4.10a). For three-regular graphs, this reduces to somewhere around 30% (Fig. 4.10b). Using the EMG distribution (i.e. Eq. 4.22) provides little improvement on the estimate for the three-regular graphs (Fig. 4.11b). This is unsurprising given how close to Gaussian the DOS is for these problems. For binomial graphs the improvement is more substantial, with typical errors being around 20%. As n is increased, we expect that the DOS will be better modelled by a continuous DOS, hence we would expect the error in β to be improved.

By assigning the quantum evolution a temperature, the challenge of understanding a dynamical problem has been mapped to a static problem. So far in this section, it has been shown how to reasonably estimate the associated temperature. If the temperature is too high, then the associated thermal state can be efficiently classically approximated. Results by Crosson *et al.* [145] suggest that for values of $\beta \leq 0.1$ a classical computer could simulate the associated thermal state efficiently for a three-regular graph, suggesting that for a CTQW to provide an advantage, it must operate outside this regime.

Estimating the diagonal entropy

Statistical mechanics provides a clear prescription for writing down an entropy given a partition function. This is given by

$$S^{SM} = \beta \langle H_{QW} \rangle + \ln \mathcal{Z}. \quad (4.23)$$

Using the modelled DOS, this entropy can be estimated. For the Gaussian distribution, Eq. 4.23 reduces to:

$$S^{\text{norm}} = n \ln 2 - \beta_{\text{est}} n + \frac{\sigma^2 \beta_{\text{est}}^2}{2}, \quad (4.24)$$

and for the EMG:

$$S^{EMG} = n \ln 2 + \beta_{EM} (\Delta - n) + \frac{\beta_{EM}^2}{2} (\sigma^2 - \Delta^2) - \ln (1 + \beta_{EM} \Delta). \quad (4.25)$$

The question is: does this estimated entropy reasonably estimate some measure of entropy in the system that is of interest? Although the ETH cannot be straightforwardly applied here, as the entropy is not a local observable, the ETH, and entanglement entropy have been shown to be linked in previous works [146, 147, 148]. Another likely candidate is the diagonal entropy of the system, since it is the on-diagonal behaviour of the CTQWs that is being captured through the DOS approximation.

Fig. 4.12 shows the diagonal and entanglement entropy averaged over a hundred 14-qubit binomial graphs as a function of time. The entanglement entropy, shown in blue, is calculated by tracing out half of the qubits, which are randomly selected, and calculating the von Neumann entropy of the reduced density operator. As is clear from the figure, the entanglement entropy is approximately constant. The dotted red line shows $S^{SM}/2$ assuming a Gibbs state, with the true DOS and the temperature fixed according to the energy. The dashed purple line shows $S^{\text{norm}}/2$ and the dashed green line $S^{EMG}/2$. The dashed pink line shows the diagonal entropy divided by 2. These have been scaled under the assumption that entropy is extensive. The prediction using a Gibbs state

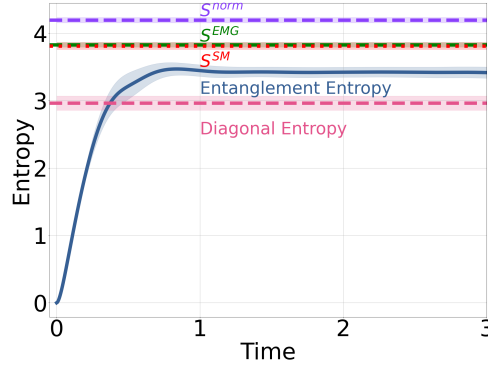


Figure 4.12: The entanglement entropy averaged over a hundred 14-qubit binomial instances (solid blue line). The dashed purple line shows $S^{\text{norm}}/2$, the dashed green line $S^{\text{EMG}}/2$, and the dotted red line shows $S^{\text{SM}}/2$. The shaded regions show a single standard deviation.

(red dotted line) overestimates the diagonal entropy (pink dashed line), this is not surprising as the Gibbs state is the maximum entropy distribution. The models for the DOS, particularly the EMG, provides a good estimate of S^{SM} . This is further evidence that the EMG is a good model for the DOS. The prediction from a Gibbs state overestimates the entanglement entropy too, but appears to provide a reasonable estimate.

As mentioned in Chapter 2, the ideal final state will be a low-entanglement state. CTQWs lack any mechanism to dissipate entanglement, as shown by Fig. 4.12.

4.6 Estimating $\overline{\langle H_p \rangle}$

In Sec. 4.5 it was shown how the temperatures associated with a CTQW could be found. By assuming the system is well modelled by a Gibbs distribution, it follows that

$$\overline{\langle H_p \rangle} = -\frac{1}{\beta} \frac{\partial \ln \mathcal{Z}}{\partial \gamma}, \quad (4.26)$$

holds true. Using Eq. 4.26 removes the need for matrix exponentiation. The derivation of this equation starts by considering the function

$$\Omega(\beta) = e^{\beta H_{QW}} \partial_\gamma e^{-\beta H_{QW}}, \quad (4.27)$$

which can be differentiated with respect to β , to get:

$$\begin{aligned} \partial_\beta \Omega &= e^{\beta H_{QW}} H_{QW} \partial_\gamma e^{-\beta H_{QW}} - e^{\beta H_{QW}} \partial_\gamma (H_{QW} e^{-\beta H_{QW}}) \\ &= -e^{\beta H_{QW}} \partial_\gamma (H_{QW}) e^{-\beta H_{QW}} \\ &= -e^{\beta H_{QW}} H_p e^{-\beta H_{QW}}. \end{aligned} \quad (4.28)$$

Now through successive differentiation or by simply applying the well-known result for writing the right-hand-side in terms of nested commutators [149] it follows that:

$$\Omega = -\beta H_p - \frac{\beta^2}{2} [H_{QW}, H_p] - \frac{\beta^3}{3!} [H_{QW}, [H_{QW}, H_p]] + \dots \quad (4.29)$$

where $[\cdot, \cdot]$ denotes the commutator. Tidying this expression up using the notation:

$$[H_{QW}^{(k)}, H_p] = \begin{cases} H_p & \text{if } k = 0 \\ \underbrace{[H_{QW}, \dots [H_{QW}, [H_{QW}, H_p]] \dots]}_{k \text{ times}} & \text{otherwise.} \end{cases}$$

gives:

$$\Omega = - \sum_{k=1}^{\infty} \frac{\beta^k}{k!} [H_{QW}^{(k-1)}, H_p]. \quad (4.30)$$

Acting on both sides with $e^{-\beta H_{QW}}$ and taking the trace gives:

$$\text{Tr} \left\{ \partial_\gamma e^{-\beta H_{QW}} \right\} = - \sum_{k=1}^{\infty} \frac{\beta^k}{k!} \text{Tr} \left\{ [H_{QW}^{(k-1)}, H_p] e^{-\beta H_{QW}} \right\}. \quad (4.31)$$

Evaluating the terms in the sum for which $k > 1$:

$$\begin{aligned}
 & \text{Tr} \left\{ \left[H_{QW}^{(k-1)}, H_p \right] e^{-\beta H_{QW}} \right\} \\
 &= \sum_j \langle E_j | \left[H_{QW}^{(k-1)}, H_p \right] e^{-\beta H_{QW}} | E_j \rangle \\
 &= \sum_j \langle E_j | \left[H_{QW}, \left[H_{QW}^{(k-2)}, H_p \right] \right] | E_j \rangle e^{-\beta E_j} \\
 &= \sum_j E_j \langle E_j | \left(\left[H_{QW}^{(k-2)}, H_p \right] - \left[H_{QW}^{(k-2)}, H_p \right] \right) | E_j \rangle e^{-\beta E_j} \\
 &= 0,
 \end{aligned}$$

where $|E_j\rangle$ denotes the eigenvectors of H_{QW} . Therefore, once it is assumed that the state is well modelled by a Gibbs distribution, it follows that:

$$\text{Tr} \left\{ \partial_\gamma e^{-\beta H_{QW}} \right\} = -\beta \text{Tr} \left\{ H_p e^{-\beta H_{QW}} \right\}. \quad (4.32)$$

Commuting through the trace with the partial derivative and dividing both sides by the partition function gives Eq. 4.26.

For the Gaussian approximation, Eq. 4.26 gives:

$$\overline{\langle H_p \rangle}_{\text{est}} = -\frac{\gamma n \kappa_2}{n + \gamma^2 \kappa_2}, \quad (4.33)$$

which is optimised by

$$\gamma_{\text{est}}^{\text{opt}} = \sqrt{\frac{n}{\kappa_2}}. \quad (4.34)$$

Note that this is the same result as the heuristic of maximising the dynamics (i.e. Eq. 4.4). The same approach can be taken for the EMG model to give:

$$\overline{\langle H_p \rangle}_{EM} = \frac{(3\kappa_3)^{1/3}}{1 + \beta_{EM} \Delta} - (3\kappa_3)^{1/3} - \left(\kappa_2 - (3\kappa_3)^{2/3} \right) \beta_{EM} \gamma. \quad (4.35)$$

Fig. 4.13 compares $\overline{\langle H_p \rangle}$ to the predictions from the Gaussian and EMG

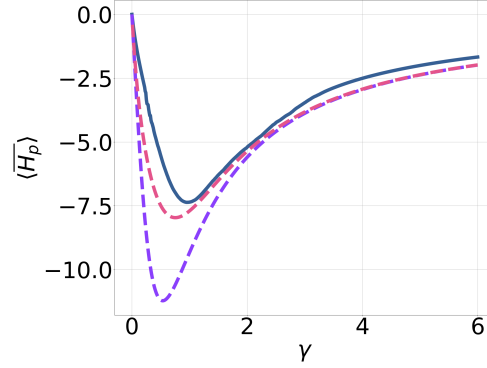


Figure 4.13: The solid blue line shows $\overline{\langle H_p \rangle}$ for a single instance of a 12-qubit binomial graph. The dashed purple (pink) line shows the prediction from Eq. 4.26 assuming a Gaussian (EMG) DOS.

predictions for a single 12-qubit binomial graph. Though both overestimate the performance of the CTQW, both provide reasonable estimates on the optimal γ . The EMG clearly provides a better model for $\overline{\langle H_p \rangle}$ than the Gaussian model. Indeed, it is remarkably close to the direct numerical calculation of $\overline{\langle H_p \rangle}$. Note that although the thermal model appears to provide good estimates for $\overline{\langle H_p \rangle}$ for large and small γ , in this regime $\langle H_p(t) \rangle$ will not necessarily display steady-state behaviour. However, this regime is unlikely to be of practical interest to CTQWs.

For the performance of $\gamma_{\text{est}}^{\text{opt}}$ the reader is referred to Sec. 4.2. For the EMG distribution, there are two important quantities to examine:

- How much does the performance of $\overline{\langle H_p \rangle}$ change between the optimal γ and the γ predicted by the EMG DOS?
- What is the difference between the prediction of $\overline{\langle H_p \rangle}$ and the true value?

The first of these is addressed for small problem sizes in Fig. 4.14. The performance is improved over γ_{heur} from simply balancing the drive and problem Hamiltonians, especially for binomial graphs. Hence, a good heuristic method for finding γ for CTQWs applied to Max-cut has been found.

Fig. 4.15 shows the error in the performance between the true value of

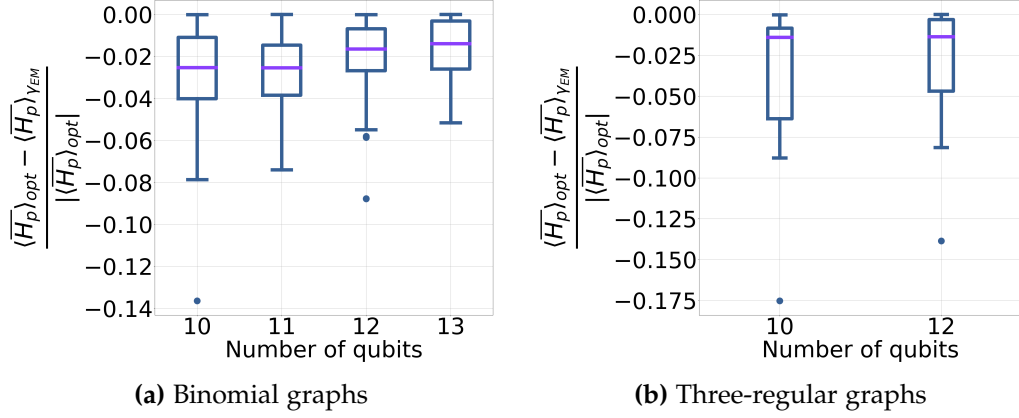


Figure 4.14: The difference in $\overline{\langle H_p \rangle}$ between the optimal choice of γ and the prediction from the EMG DOS (γ_{EM}), normalised by the absolute value of $\langle H_p \rangle$.

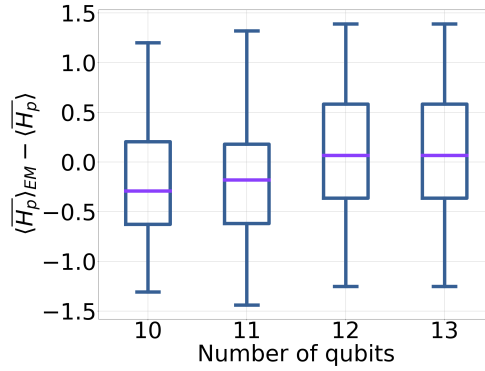


Figure 4.15: The difference between the prediction of $\langle H_p \rangle$ with an EMG DOS using Eq. 4.26 (i.e. $\langle H_p \rangle_{EM}$) and the true value. For each instance, the value of γ has been chosen such that $\langle H_p \rangle_{EM}$ is minimised. Each bar shows one hundred binomial graph instances.

$\overline{\langle H_p \rangle}$ and the predicted value from the EMG approximation, for the γ predicted to give the best possible performance from the EMG DOS. The error is relatively small for these small problem sizes. Given the tractability of Eq. 4.26, and evidence of small errors, it is possible to estimate the performance for larger problem sizes. This is shown in Fig. 4.16.

By treating closed-system CTQWs as thermalising systems, it has been possible to provide useful heuristic choices for optimising γ and to make practical predictions for $\overline{\langle H_p \rangle}$.

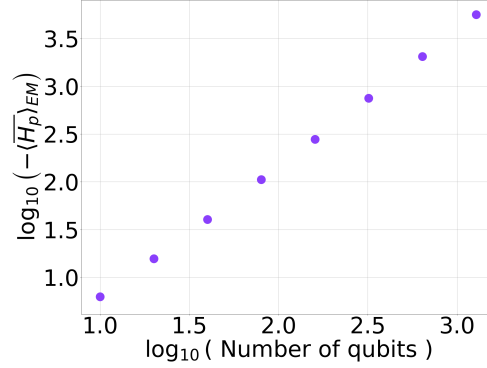


Figure 4.16: Estimated scaling of the CTQW for binomial graphs, assuming thermalisation and an EMG DOS. Each data point shows the median performance of one hundred binomial graph instances.

4.7 The performance of CTQWs for Max-cut

In this section, the value of $\overline{\langle H_p \rangle}$ is numerically evaluated at the γ predicted by the EMG DOS, for the problem instances considered in Sec. 4.2. This is then mapped to a time-averaged mean cut-value i.e. $(\kappa_2 - \overline{\langle H_p \rangle})/2$. This is then compared to the maximum cut, i.e. $(\kappa_2 - E_0^{(P)})/2$. This choice is to make easier comparison to existing literature. For example, it is known that QAOA $p = 1$ lower bounds $r = (\kappa_2 - \overline{\langle H_p \rangle})/(\kappa_2 - \overline{\langle H_p \rangle})$ by 0.692 on three-regular graphs [14]. While QAOA $p = 2$ lower bounds r by 0.7559 on three-regular graphs [150]. The classical Goemans-Williamson algorithm [151] lower bounds r by 0.8786, and by 0.9326 on three-regular graphs [152]. The results for the CTQW can be seen in Fig. 4.17. For the small problem sizes considered, the performance seems largely independent of system size. This perhaps is unsurprising for a high temperature Gibbs state with a local Hamiltonian [153, 154]. Since both $\overline{\langle H_p \rangle}$ and $E_0^{(P)}$ are extensive, this is to be expected. The CTQW appears to give better results for binomial graphs, this may be that the EMG is better capturing the frustration in the problem. If the frustration is entering the problem through loops with length greater than three, this is not captured by the ansatz.

The scaling of r being approximately independent of n does not mean

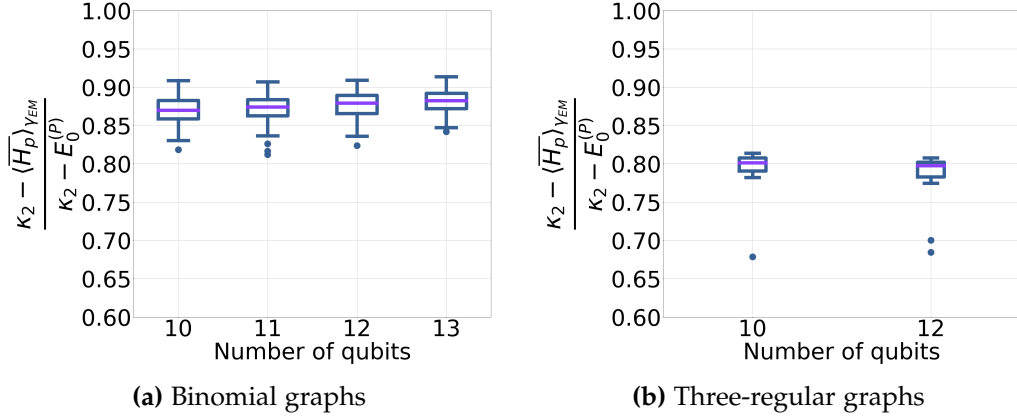


Figure 4.17: The normalised performance in terms of the time averaged mean cut value (i.e., $(\kappa_2 - \langle H_p \rangle)/2$) of the CTQW for the optimal γ predicted by the EMG DOS compared to the maximum cut (i.e. $(\kappa_2 - E_0^{(p)})/2$) on the graph

that the ground-state probability is constant with problem-size. Given that Max-cut is a hard problem, as discussed in Sec. 2.4, it is reasonable to suspect that the ground-state probability is decreasing exponentially. The exponential decrease in ground-state probability has been seen in other problems that CTQWs have been applied [22].

In practice, one is perhaps less interested in how the average of the energy distribution of $\langle H_p \rangle$ compares with the absolute minimum, particularly if considering CTQWs as an exact solver. In such a case, one might be more interested in the ground-state probability or time-to-solution.

4.8 Conclusion

Throughout this chapter, we have attempted to understand the performance of CTQWs outside what is classically simulable for non-integrable models. For the steady-state, it was conjectured that the system thermalises. This means that, despite the state-vector consisting of 2^n complex amplitudes, for a given γ the value of $\langle H_p \rangle$ will depend on one real number, the energy. This provides insight into the computational mechanisms involved in CTQWs.

Assuming thermalisation, classical statistical physics provides an alternative route to understanding CTQWs for Max-cut on a broad range of graphs within a closed-system setting. By associating the unitary dynamics with a temperature, it provides an alternative route to achieve the same value of $\overline{\langle H_p \rangle}$, either through dissipative dynamics or classical simulation. Here we have introduced an EMG distribution to account for some frustration in the model to make analytic predictions. This model provides an estimate of the energy distribution H_{QW} but says nothing about the strings involved (including the ground-state solution) or the energy distribution associated with just H_p . By exploiting this model, it was possible to find reasonable estimates for the optimal choice of γ , that utilised properties of the underlying graph. Importantly, predictions far away from what it is easy to directly simulate were made.

Chapter 5

Thermalisation timescales

5.1 Introduction

Understanding only the long-time behaviour of CTQWs is insufficient. If, for example, the time taken to thermalise grows exponentially, then the long-time behaviour is of no practical interest for optimisation. This section investigates the timescale associated with CTQWs, again focusing on Max-cut with a transverse-field driver Hamiltonian. Sec. 5.2 takes a geometrical approach, evaluating the curvature to give a timescale. Sec. 5.3 analytically investigates the very short-time limit of CTQWs for Max-cut. Sec. 5.4 explores a simple Ansatz for the time evolution of CTQWs in order to extract a timescale.

5.2 A geometrical description of the dynamics

Classically, a one-dimensional curve in Euclidean space can be described in terms of geometrical properties — such as curvature and torsion. Laba *et al.* [25] extended these ideas to Schrödinger evolution under a time-independent Hamiltonian. Curvature measures how far the evolution of a state-vector deviates from a geodesic. If the curvature is zero, then the evolution is completely described by a great circle on a Bloch sphere. The curvature for a time-independent Hamiltonian H is given by:

$$\mathcal{C} = \langle \Delta H^4 \rangle - \langle \Delta H^2 \rangle^2, \quad (5.1)$$

where $\Delta H = H - \langle H \rangle$. Since it depends on only expectation values of powers of the Hamiltonian, it is conserved over the evolution. The curvature can be seen as a measure of inefficiency. The torsion is given by:

$$\mathcal{T} = \mathcal{C} - \frac{\langle \Delta H^3 \rangle^2}{\langle \Delta H^2 \rangle}. \quad (5.2)$$

The torsion measures how much the wave-function deviates from a two-dimensional space in a time Δt . The error from deviating from the two-dimensional subspace in this time is given by [136]:

$$\varepsilon_{2D} = \mathcal{T} \Delta t^4. \quad (5.3)$$

with $\varepsilon_{2D} = 0$ corresponding to remaining in the two-dimensional subspace and $\varepsilon_{2D} = 1$ corresponding to departing the subspace. If $\varepsilon_{2D} \ll 1$, then the two-dimensional approximation holds well. That is, for times $t \ll \mathcal{T}^{-1/4}$, the evolution can be approximated by a two-level system. Therefore, $\mathcal{T}^{-1/4}$ sets a timescale for CTQWs.

Gnatenko *et al.* proposed using a quantum computer to measure torsion and curvature to calculate certain graph properties. This section inverts these ideas to analytically calculate the curvature and torsion of CTQWs for Max-cut. Calculating the torsion and curvature reduces to evaluating $\langle \Delta H_{QW}^j \rangle$, for $j = 2, 3, 4$. Since these are all conserved quantities, the expectation value can be taken with respect to the initial state, $|+\rangle$. The first step is to find $\langle H_{QW} \rangle$:

$$\begin{aligned} \langle H_{QW} \rangle &= \langle + | H_d + \gamma H_p | + \rangle \\ &= -n, \end{aligned}$$

since each $Z_i Z_j$ in H_p will flip qubits i and j to the minus state. Consequently, only H_d contributes to this expectation value.

Consider now, $\langle \Delta H_{QW}^2 \rangle = \langle H_{QW}^2 \rangle - \langle H_{QW} \rangle^2$:

$$\begin{aligned}
\langle \Delta H_{QW}^2 \rangle &= \langle H_d^2 \rangle + \gamma^2 \langle H_p^2 \rangle + \gamma \langle H_p H_d + H_d H_p \rangle - \langle H \rangle^2 \\
&= n^2 + \gamma^2 \kappa_2 + 0 + 0 - n^2 \\
&= \gamma^2 \kappa_2,
\end{aligned}$$

where $\kappa_2 = |E|$ is the number of edges in the graph. It was pointed out by Gnatenko *et al.* [136], that $\langle H_p^2 \rangle$ is equal to the number of edges in the graph. Writing out this term explicitly gives:

$$\langle H_p^2 \rangle = \sum_{i_1, j_1, i_2, j_2} \langle + | Z_{i_1} Z_{j_1} Z_{i_2} Z_{j_2} | + \rangle,$$

where the sum is over the ordered pairs (i_k, j_k) , for $k = 1, 2$. The only possible way a term in the sum can be non-zero is if $i_1 = i_2$ and $j_1 = j_2$. Therefore, the number of non-zero terms corresponds to the number of edges in the graph. The rest of the terms in $\langle \Delta H^2 \rangle$ are relatively trivial to calculate.

Extending the above logic, Gnatenko *et al.* [136] demonstrated that:

$$\langle H_p^3 \rangle = 6\kappa_3 \tag{5.4}$$

where κ_3 is the number of triangles in the graph, and

$$\langle H_p^4 \rangle = \kappa_2 + 3\kappa_2 (\kappa_2 - 1) + 24\kappa_4, \tag{5.5}$$

where κ_4 is the number of squares in the graph.

Using the above results, $\langle \Delta H_{QW}^3 \rangle$ and $\langle \Delta H_{QW}^4 \rangle$ can be calculated. Starting with $\langle \Delta H_{QW}^3 \rangle$:

$$\langle \Delta H_{QW}^3 \rangle = \langle H_{QW}^3 \rangle - 3\langle H_{QW}^2 \rangle \langle H_{QW} \rangle + 2\langle H_{QW} \rangle^3. \tag{5.6}$$

Calculating $\langle H_{QW}^3 \rangle$ gives:

$$\begin{aligned}\langle H_{QW}^3 \rangle &= \langle (H_d + \gamma H_p)^3 \rangle \\ &= \langle H_d^3 + \gamma^2 H_d H_p^2 + \gamma^2 H_p^2 H_d + \gamma^3 H_p^3 + \gamma^2 H_p H_d H_p \rangle \\ &= -n^3 - 2n\gamma^2 \kappa_2 + 6\gamma^3 \kappa_3 + \gamma^2 \langle H_p H_d H_p \rangle,\end{aligned}$$

where terms that are trivially zero have been neglected and Eq. 5.4 has been applied. To find $\langle \Delta H_{QW}^3 \rangle$ it just remains to calculate $\langle H_p H_d H_p \rangle$:

$$\langle H_p H_d H_p \rangle = - \sum_{i_1, j_1, i_2, j_2, k} \langle + | Z_{i_1} Z_{j_2} X_k Z_{i_2} Z_{j_2} | + \rangle. \quad (5.7)$$

The X_k in each term of the sum is not going to change which terms are non-zero. It will only introduce a minus sign if k is equal to i_2 or j_2 . Hence, $\langle H_p H_d H_p \rangle$ is equal to:

$$\begin{aligned}&= - \sum_{i_2, j_2, k} \langle + | Z_{i_2} Z_{j_2} X_k Z_{i_2} Z_{j_2} | + \rangle \\ &= - \sum_{\substack{i_2, j_2 \\ k \neq i_2, j_2}} \langle + | Z_{i_2} Z_{j_2} Z_{i_2} Z_{j_2} | + \rangle + \sum_{j_2} \langle + | Z_k Z_{j_2} Z_k Z_{j_2} | + \rangle \\ &\quad + \sum_{i_2} \langle + | Z_{i_2} Z_k Z_{i_2} Z_k | + \rangle \\ &= -(n-2)\kappa_2 + \kappa_2 + \kappa_2 \\ &= -(n-4)\kappa_2.\end{aligned}$$

Alternatively, arguing as follows: in the absence of the X_k term, Eq. 5.7 would give $-n\kappa_2$. Accounting for the edge cases where $k = i_2$ or $k = j_2$ gives:

$$\langle H_p H_d H_p \rangle = -(n-4)\kappa_2. \quad (5.8)$$

Combining all of the above, the result is:

$$\langle \Delta H_{QW}^3 \rangle = 2\gamma^2 (2\kappa_2 + 3\gamma\kappa_3). \quad (5.9)$$

To find \mathcal{C} and \mathcal{T} , it remains to find $\langle \Delta H_{QW}^4 \rangle$.

$$\begin{aligned} \langle \Delta H_{QW}^4 \rangle &= \langle H_{QW}^4 \rangle - 4\langle H_{QW} \rangle \langle H_{QW}^3 \rangle + 6\langle H_{QW} \rangle^2 \langle H_{QW}^2 \rangle \\ &\quad - 4\langle H_{QW} \rangle^3 \langle H_{QW} \rangle + \langle H_{QW} \rangle^4 \end{aligned}$$

Evaluating $\langle H_{QW}^4 \rangle$ gives:

$$\begin{aligned} \langle H_{QW}^4 \rangle &= \langle H_d^4 + \gamma^2 (H_d^2 H_p^2 + H_d H_p H_d H_p + H_d H_p^2 H_d + H_p H_d^2 H_p \\ &\quad + H_p H_d H_p H_d + H_p^2 H_d^2) + \gamma^3 (H_d H_p^3 + H_p^2 H_d H_p + H_p^3 H_d + H_p H_d H_p^2) \\ &\quad + \gamma^4 H_p^4 \rangle, \end{aligned}$$

where the terms that evaluate to zero have been neglected. Most of the terms in this expansion have already been evaluated, or are simple to evaluate. This leaves $\langle H_p H_d^2 H_p \rangle$, $\langle H_p H_d H_p^2 \rangle$, $\langle H_p^2 H_d H_p \rangle$, and $\langle H_p^4 \rangle$, the last of which can be evaluated with Eq. 5.5. Turning now to $\langle H_p H_d^2 H_p \rangle$:

$$\langle H_p H_d^2 H_p \rangle = \sum_{i_1, j_1, i_2, j_2, k_1, k_2} \langle + | Z_{i_1} Z_{j_1} X_{k_1} X_{k_2} Z_{i_2} Z_{j_2} | + \rangle.$$

Again, the presence of X 's in the sum will not change which terms evaluate to be non-zero. Fixing the edge (i_2, j_2) , if either one of the X terms coincides with this edge and the other X term does not, it contributes -1 to the sum; there are $4(n-2)$ ways of this happening. If both X terms do not coincide with this edge, then the term contributes 1; there are $(n-2)^2$ ways of this happening. Finally, if both edges coincide, this edge contributes 1; there are four ways of this happening. The result is:

$$\langle H_p H_d^2 H_p \rangle = (n-4)^2 \kappa_2.$$

Thus, it remains to evaluate $\langle H_p H_d H_p^2 \rangle$ and $\langle H_p^2 H_d H_p \rangle$. Note that:

$$\langle H_p H_d H_p^2 \rangle^* = \langle H_p^2 H_d H_p \rangle.$$

To evaluate $\langle H_p H_d H_p^2 \rangle$, consider $\langle H_p^3 \rangle$. As before, the introduction of H_d is not going to change which terms in the sum are non-zero, if we were to write out a similar sum for this term, as in Eq. 5.7. Therefore, $\langle H_p H_d H_p^2 \rangle$ is going to depend on the number of triangles in the graph. The introduction of H_d is going to result in sign flips to some terms. The result is

$$\langle H_p H_d H_p^2 \rangle = -6\kappa_3(n-4). \quad (5.10)$$

Finally, assembling all the expectation values gives

$$\langle \Delta H_{QW}^4 \rangle = \gamma^2 \left[-2\kappa_2(\gamma^2 - 8) + 3\gamma^2 \kappa_2^2 + 24\gamma(2\kappa_3 + \gamma\kappa_4) \right]. \quad (5.11)$$

The resulting expression for the curvature is:

$$\mathcal{C} = 2\gamma^2 \left[-\kappa_2(\gamma^2 - 8) + \gamma^2 \kappa_2^2 + 12\gamma(2\kappa_3 + \gamma\kappa_4) \right]. \quad (5.12)$$

and for the torsion:

$$\mathcal{T} = 2\gamma^4 \left[\kappa_2(\kappa_2 - 1) - \frac{18\kappa_3^2}{\kappa_2} + 12\kappa_4 \right]. \quad (5.13)$$

Note that the torsion is zero for a two-qubit system, with one edge, as expected from the spin-flip symmetry in the problem. The torsion is also zero for a ring of three qubits. It is clear that the torsion is reduced by the number of triangles in the graph. This suggests that the CTQW explores the solution space slower for graphs with many triangles.

5.3 The very short-time limit

In the previous section, the torsion of CTQWs for Max-cut was calculated. As mentioned, in the time interval $\delta t \ll \mathcal{T}^{-1/4}$, the evolution of a CTQW can be approximated by a two-level system. The initial-state of a CTQW is known, therefore it is possible to analytically calculate the exact evolution of a CTQW in the interval $t \ll \mathcal{T}^{-1/4}$. The rest of this section carries out this calculation.

At $t = 0$ the state of the system is $|\psi(t=0)\rangle = |+\rangle$. Integrating the system forward some infinitesimal time δt , gives $|\psi(t=\delta t)\rangle = (I - iH_{QW}\delta t)|+\rangle$. Applying the Gram-Schmidt procedure to the two vectors, $|\psi(t=0)\rangle$ and $|\psi(t=\delta t)\rangle$, gives the orthonormal basis for the subspace:

$$|e_0\rangle = |+\rangle \quad (5.14)$$

$$|e_1\rangle = \Delta H_{QW} |+\rangle / \sqrt{\langle \Delta H_{QW}^2 \rangle}, \quad (5.15)$$

where the expectation is with respect to the $|+\rangle$ state. Writing out H in this subspace is:

$$H_{QW}^{(2D)} = \begin{pmatrix} \langle H_{QW} \rangle & \frac{\langle H_{QW} \Delta H_{QW} \rangle}{\sqrt{\langle \Delta H_{QW}^2 \rangle}} \\ \frac{\langle \Delta H_{QW} H_{QW} \rangle}{\sqrt{\langle \Delta H_{QW}^2 \rangle}} & \frac{\langle \Delta H_{QW} H_{QW} \Delta H_{QW} \rangle}{\langle \Delta H_{QW}^2 \rangle} \end{pmatrix}, \quad (5.16)$$

which evaluates to

$$H_{QW}^{(2D)} = \begin{pmatrix} -n & \sqrt{\gamma^2 \kappa_2} \\ \sqrt{\gamma^2 \kappa_2} & 2 \left(2 + \frac{3\gamma \kappa_3}{\kappa_2} \right) - n \end{pmatrix} \quad (5.17)$$

$$(5.18)$$

which equals, up to terms proportional to the identity,

$$H_{QW}^{(2D)} = \sqrt{\gamma^2 \kappa_2} X - \left(2 + \frac{3\gamma \kappa_3}{\kappa_2} \right) Z. \quad (5.19)$$

The problem Hamiltonian, within the same space is:

$$H_p^{(2D)} = \begin{pmatrix} \langle H_p \rangle & \frac{\langle H_p \Delta H \rangle}{\sqrt{\langle \Delta H^2 \rangle}} \\ \frac{\langle \Delta H H_p \rangle}{\sqrt{\langle \Delta H^2 \rangle}} & \frac{\langle \Delta H H_p \Delta H \rangle}{\langle \Delta H^2 \rangle} \end{pmatrix}, \quad (5.20)$$

evaluating this gives:

$$H_p^{(2D)} = \begin{pmatrix} 0 & \text{sgn}(\gamma) \sqrt{\kappa_2} \\ \text{sgn}(\gamma) \sqrt{\kappa_2} & \frac{6\kappa_3}{\kappa_2} \end{pmatrix}, \quad (5.21)$$

where $\text{sgn}(\gamma)$ denotes the sign of γ . Having now explicitly evaluated the relevant operators in this subspace, it remains to calculate the expectation of the problem Hamiltonian, which is:

$$\langle H_p^{(2D)}(t) \rangle = -4\gamma\kappa_2 \frac{\sin^2 \omega t}{\omega^2}, \quad (5.22)$$

where

$$\omega^2 = \gamma^2 \kappa_2 + \frac{(2\kappa_2 + 3\gamma\kappa_3)^2}{\kappa_2^2}. \quad (5.23)$$

As we can see, for very short times, $\langle H_p \rangle$ depends very little on the properties of the underlying Max-cut graph, depending primarily on the number of edges in the graph. The frequency term ω does depend on the number of triangles but for large problem sizes it is reasonable to expect $\omega^2 \approx \gamma^2 \kappa_2$. Physically it is reasonable that, at short times, the CTQW only sees triangles and edges. Longer evolution would result in the approach seeing larger loops in the graph.

From a computational point of view this suggests that any useful short-time sampling must not be on a timescale much smaller than $\mathcal{T}^{-\frac{1}{4}}$ since you would like the CTQW to see the whole graph. A rough estimate for a CTQW to see the whole structure of the graph might be $l\mathcal{T}^{-\frac{1}{4}}$, where l is the length of the largest loop in the problem graph. This is bounded from above by $n\mathcal{T}^{-\frac{1}{4}}$.

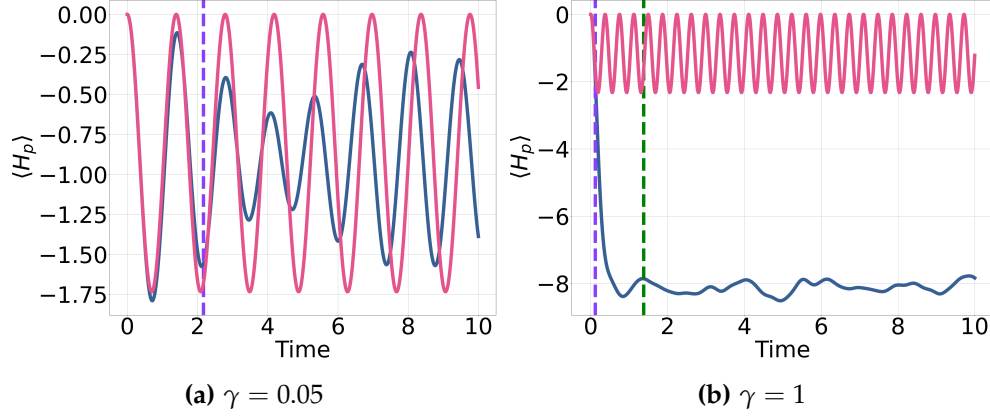


Figure 5.1: The performance of a CTQW on two different 12-qubit binomial graphs with different values of γ . The blue line shows the result of direct integration of the Schrödinger equation. The pink line shows the result of the two-dimensional approximation (i.e, Eq. 5.22). The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$. The dashed green line in (5.1b) shows the location of $n\mathcal{T}^{-\frac{1}{4}}$, with $n = 12$. The values of γ have been chosen to illustrate different behaviour and do not correspond to the optimal γ for the graph.

In Fig. 5.1a, $\langle H_p^{(2D)}(t) \rangle$ (i.e. Eq. 5.22) is compared to numerical simulation with a 12-qubit binomial graph with $\gamma = 0.05$. The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$. For short times, there is good quantitative agreement between the numerical simulation (the blue line) and the two-level prediction (the pink line). At longer times, there is reasonable qualitative agreement, with the two-dimensional approximation capturing the oscillatory nature.

Examining a different 12-qubit binomial graph with $\gamma = 1$ gives Fig. 5.1b. Again, we can see good agreement between the numerical simulation (the blue line) and the two-level prediction (the pink line) for short times. The dashed purple line corresponds to $\mathcal{T}^{-\frac{1}{4}}$ with the dashed green line corresponding to $n\mathcal{T}^{-\frac{1}{4}}$. Here, the two-level approximation provides poor qualitative insight into the CTQW outside the short-time limit, with the CTQW approaching an approximate steady state. Noticeably, by $n\mathcal{T}^{-\frac{1}{4}}$, the system has settled into the steady state.

5.4 An Ansatz for the Schrödinger Evolution

In Chapter 3 it was discussed that thermalisation in a closed-system is as a result of dephasing. In this section, inspired by the off-diagonal elements of H_p , we take a Gaussian Ansatz for the Schrödinger evolution. Using this, it is possible to extract a timescale that better captures the thermalisation timescale.

The difference between $\langle H_p(t) \rangle$ and $\overline{\langle H_p \rangle}$ is given by

$$\langle \Delta H_p(t) \rangle = \langle H_p(t) \rangle - \overline{\langle H_p \rangle} \quad (5.24)$$

$$= \sum_{E_m \neq E_n} e^{-i(E_m - E_n)t} \langle E_m | A | E_n \rangle \langle E_n | \rho_0 | E_m \rangle, \quad (5.25)$$

where $|E_n\rangle$ is an eigenstate of H_{QW} with eigenvalue E_n and initial-state ρ_0 . Let $\omega_{m,n} = E_m - E_n$ and $\nu_{m,n} = \langle E_m | H_p | E_n \rangle \langle E_n | \rho_0 | E_m \rangle$, then:

$$\langle \Delta H_p(t) \rangle = \sum_{\omega_{m,n}} e^{-i\omega_{m,n}t} \nu_{m,n} \quad (5.26)$$

$$\approx \int_{-\infty}^{\infty} e^{-i\omega t} \nu(\omega) \Omega(\omega) d\omega. \quad (5.27)$$

The above approximates the sum in Eq. 5.26 as an integral in Eq. 5.27, where $\nu(\omega)$ is the coarse-grained version of $\nu_{m,n}$ and $\Omega(\omega)$ the density of gaps. From Eq. 5.27 it is clear that $\langle \Delta H_p(t) \rangle$ is associated with the Fourier transform of the off-diagonal elements of H_p in the energy eigenbasis. It follows that the more spread the off-diagonal matrix elements are, the faster the thermalisation.

Fig. 5.2 shows the distribution of $\nu_{m,n} = \langle E_m | H_p | E_n \rangle \langle E_n | \rho_0 | E_m \rangle$ for a CTQW Hamiltonian based on an 11-qubit binomial graph. The matrix values are peaked around 0. Approximately modelling the matrix values as a Gaussian (i.e. taking $\nu(\omega)\Omega(\omega)$ to be a Gaussian) gives the following Ansatz:

$$\langle \Delta H_p(t) \rangle = B e^{-t^2/\tau^2}, \quad (5.28)$$

where τ and B are parameters that need to be set. Expanding this equa-

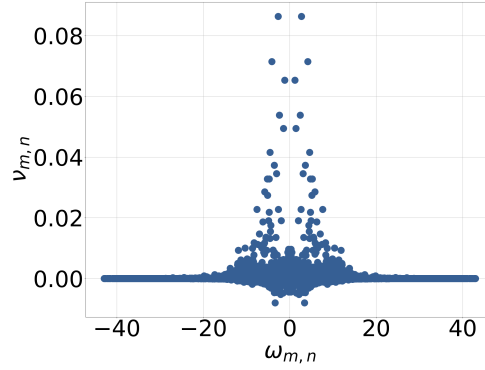


Figure 5.2: The values of $\nu_{m,n}$ for an 11-qubit binomial graph with $\gamma = 1$

tion to second order in time gives:

$$\langle \Delta H_p(t) \rangle = B \left(1 - t^2 / \tau^2 + \dots \right). \quad (5.29)$$

From the Schrödinger dynamics,

$$\begin{aligned} \langle \Delta H_p(t) \rangle &= \langle + | H_p | + \rangle - it \langle + | [H_{QW}, H_p] | + \rangle \\ &\quad - \frac{t^2}{2} \langle + | [H_{QW}, [H_{QW}, H_p]] | + \rangle + \dots - \overline{\langle H_p \rangle}. \end{aligned} \quad (5.30)$$

Simplifying the above gives

$$\langle \Delta H_p(t) \rangle = -\overline{\langle H_p \rangle} - \frac{\gamma t^2}{2} \langle + | [H_p, [H_d, H_p]] | + \rangle + \dots \quad (5.31)$$

Matching terms gives:

$$B = -\overline{\langle H_p \rangle} \quad (5.32)$$

$$\frac{1}{\tau^2} = -\frac{\gamma \langle + | [H_p, [H_d, H_p]] | + \rangle}{2\overline{\langle H_p \rangle}}. \quad (5.33)$$

Evaluating the matrix element gives:

$$\frac{1}{\tau^2} = -\frac{4\gamma\kappa_2}{\overline{\langle H_p \rangle}}. \quad (5.34)$$

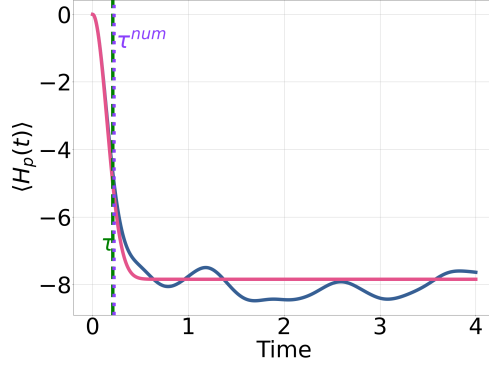


Figure 5.3: An example of a CTQW on a 12-qubit binomial graph. The solid blue line shows the Schrödinger evolution of $\langle H_p(t) \rangle$. The solid red line shows Eq. 5.35 with τ given by Eq. 5.34. The dashed green line shows the location of τ . The dotted purple line shows the location of τ extracted numerically from the Schrödinger evolution (i.e. Eq. 5.36).

where κ_2 is the number of edges in the Max-cut graph. Typically, $\gamma \overline{\langle H_p \rangle}$ will be proportional to n . For an optimised CTQW, γ will scale to balance the energy between H_d and H_p , so scales as approximately $n^{-1/2}$. Since κ_2 scales as n^2 , $\frac{1}{\tau^2}$ is expected to be largely independent of n .

Fig. 5.3 shows $\langle H_p(t) \rangle$ for a 12-qubit CTQW in blue. In this example $\gamma = 1$. The red line shows the prediction from Eq. 5.31, i.e,

$$\langle H_p \rangle_{\text{Gauss}} = \overline{\langle H_p \rangle} \left(1 - e^{-\frac{t^2}{\tau^2}} \right). \quad (5.35)$$

The timescale τ is also marked on the graph by the dashed green line. The dashed purple line shows τ^{num} extracted from the Schrödinger evolution and is defined by

$$\langle H_p(\tau^{\text{num}}) \rangle = \overline{\langle H_p \rangle} \left(1 - e^{-1} \right). \quad (5.36)$$

Visually there is good agreement between the Schrödinger evolution and Eq. 5.35. For this instance $\tau \approx 0.21$ and $\tau^{\text{num}} \approx 0.22$.

Fig. 5.4a shows τ^{num} for many instances. Fig 5.4b shows the error between τ and τ^{num} for the same instances. In each instance $\gamma = \sqrt{n/\kappa_2}$. In all the cases considered the timescale τ appears largely independent

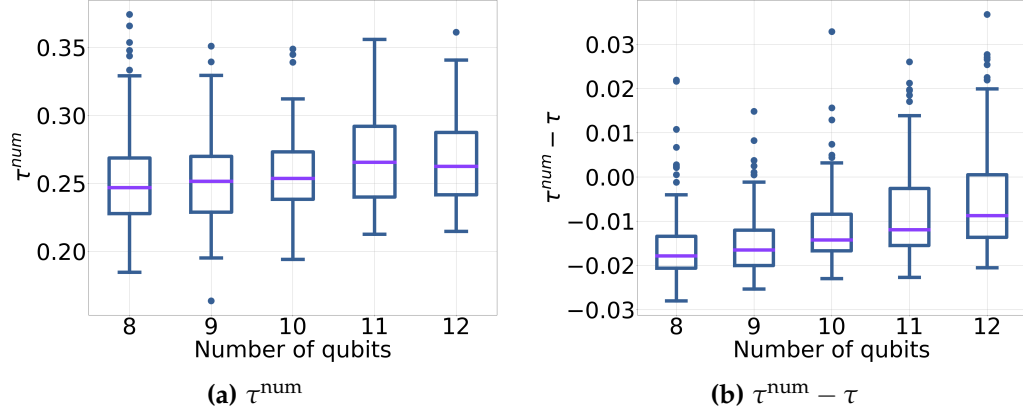


Figure 5.4: The thermalisation timescale for a closed-system CTQW on n qubits. For each instance, 100 examples are considered.

of the problem size as expected and has good agreement with τ^{num} .

So far, only small-scale numerics have been explored. To try to validate the Gaussian Ansatz (i.e, Eq. 5.35) outside this limit, we might try and see how well it predicts higher order terms in Eq. 5.30. For a CTQW, consisting of a real Hamiltonian, $\nu_{m,n}$ is real and $\nu_{m,n} = \nu_{n,m}$. Therefore, the coefficients of all the odd powers of t vanish. This is correctly predicted by the Gaussian Ansatz (or any real, even Ansatz). Therefore, the first correction in the limit of small time δt is given by:

$$\begin{aligned} \langle H_p(\delta t) \rangle - \langle H_p(\delta t) \rangle_{\text{Gauss}} &= \left(\frac{1}{4!} \langle + | [H_{QW}, [H_{QW}, [H_{QW}, [H_{QW}, H_p]]]] | + \rangle \right. \\ &\quad \left. + \frac{(4\gamma\kappa_2)^2}{2\langle H_p \rangle} \right) \delta t^4 + \mathcal{O}(\delta t^6) \end{aligned} \quad (5.37)$$

The matrix element in the above equation is given by:

$$\begin{aligned} \langle + | [H_{QW}, [H_{QW}, [H_{QW}, [H_{QW}, H_p]]]] | + \rangle &= 32\gamma \left[3\gamma (4\kappa_3 + \gamma\kappa_{2,2}) + \kappa_2 (4 + \gamma^2) \right], \end{aligned} \quad (5.38)$$

where κ_3 is the number of triangles in the Max-cut graph and $\kappa_{2,2}$ counts

the number of edges that share a node:

$$\kappa_{2,2} = \frac{1}{2} \sum_{(i,j) \in E} (\deg(i) + \deg(j) - 2) \quad (5.39)$$

where $\deg(i)$ is the degree of the i^{th} node.

To try to understand how important the error is in Eq. 5.37, to leading order the error is:

$$\varepsilon_{\text{Gauss}}(\delta t) = \frac{\langle H_p(\delta t) \rangle - \langle H_p(\delta t) \rangle_{\text{Gauss}}}{\langle H_p(\delta t) \rangle} \quad (5.40)$$

$$\approx \frac{\frac{4\gamma}{3} [3\gamma(4\kappa_3 + \gamma\kappa_{2,2}) + \kappa_2(4 + \gamma^2)] + \frac{(4\gamma\kappa_2)^2}{2\langle H_p \rangle}}{-4\gamma\kappa_2} \delta t^2 + \mathcal{O}(\delta t^4), \quad (5.41)$$

for an appropriately chosen δt . The choice of δt needs to be small enough that expanding the Schrödinger equation up to fourth order in δt is a good approximation, but an expansion up to δt^2 is not sufficient. A rough estimate of this time is captured by the torsion of the wavefunction, i.e. $\delta t = \mathcal{T}^{-1/4}$. This was calculated analytically in Sec. 5.2.

The only term that cannot be straightforwardly analytically calculated in Eq. 5.41 is $\langle H_p \rangle$. To estimate this, the EMG Ansatz for the DOS is used. This means that the resulting data is a test of both the EMG Ansatz for the DOS and the Gaussian Ansatz for the time evolution.

In Fig. 5.5 $\varepsilon_{\text{Gauss}}(\mathcal{T}^{-1/4})$ (i.e. Eq. 5.41) is plotted for various problem sizes. For each instance $\gamma = \sqrt{n/\kappa_2}$. The error is relatively small at about 2% for the problem instances considered up to 256 qubits. The error appears to be tending to zero for larger problem sizes.

In summary, Eq. 5.35 appears to be a good model for the time evolution of a CTQW for Max-cut, with the associated timescale given by Eq. 5.34. In Chapter 4 it was shown how $\langle H_p \rangle$ might be reasonably estimated, hence how τ might be estimated for a CTQW. This concludes how CTQWs for Max-cut should be set up. An Ansatz, say an EMG, is chosen for the DOS which is used to select the optimal γ . The optimised

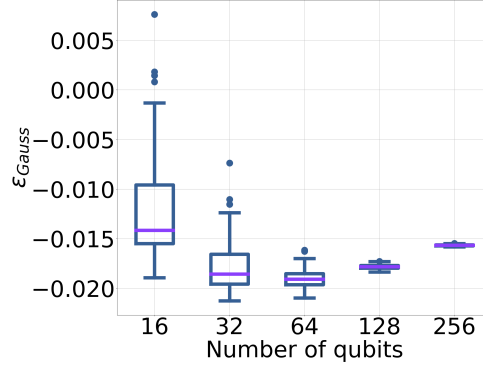


Figure 5.5: The error between the Gaussian Ansatz and the Schrödinger equation (i.e. Eq. 5.41) at $t = \mathcal{T}^{-1/4}$. At each problem size, 100 instances at Max-cut on binomial graphs are considered with $\gamma = \sqrt{n/\kappa_2}$. The value of $\langle H_p \rangle$ has been estimated using an EMG Ansatz for the DOS.

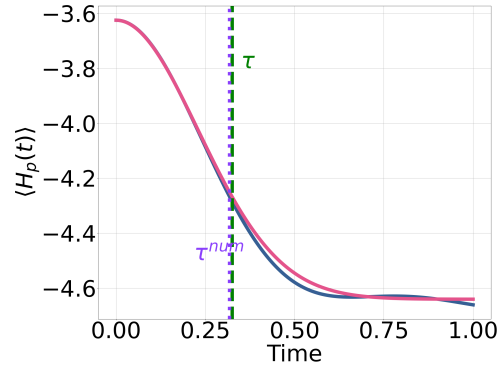


Figure 5.6: An example of a CTQW on a 10-qubit binomial graph with $\gamma_1 = 0.705$. The solid blue line shows the Schrödinger evolution of $\langle H_p(t) \rangle$. The solid red line shows Eq. 5.44 with τ given by Eq. 5.46. The dashed green line shows the location of τ . The dotted purple line shows the location of τ extracted numerically from the Schrödinger evolution (i.e. Eq. 5.36). The initial state is a Gibbs state with $\beta = 0.5$ and $\gamma_0 = 0.440$.

Hamiltonian is then prepared on the quantum device, which is allowed to run for $t \gg \tau$, say $t \approx 3\tau$, with τ estimated from the DOS too. After this time, the state is sampled. Although this work has focused on Max-cut, the process laid out is transferrable to any other Ising Hamiltonian. Given that the physics, such as thermalisation, is expected to hold for a wide range of Hamiltonians, the process should be widely applicable.

5.5 A thermal start

To extend the conclusions of the previous section, we consider the case where the initial state ρ_0 is a thermal state with inverse temperature β and Hamiltonian $H_0 = H_d + \gamma_0 H_p$, such that:

$$\rho_0 = \frac{e^{-\beta H_0}}{\text{Tr } e^{-\beta H_0}}. \quad (5.42)$$

This has connections to multi-stage quantum walks, which will be explored in Chapter 6. The CTQW Hamiltonian in this case is given by

$$H'_{QW} = H_0 + \gamma_1 H_p. \quad (5.43)$$

Again, making use of a Gaussian Ansatz for $\langle H_p(t) \rangle$,

$$\langle H_p(t) \rangle_{\text{Gauss}} = \overline{\langle H_p \rangle} + \left(\langle H_p(0) \rangle - \overline{\langle H_p \rangle} \right) e^{-t^2/\tau^2}, \quad (5.44)$$

where

$$\langle H_p(0) \rangle = \text{Tr} (H_p \rho_0) \quad (5.45)$$

$$\frac{1}{\tau^2} = \frac{(\gamma_1 - \gamma_0) \text{Tr} ([H_p, [H_d, H_p]] \rho_0)}{2 \left(\langle H_p(0) \rangle - \overline{\langle H_p \rangle} \right)}. \quad (5.46)$$

This is shown in Fig. 5.6 for a 10-qubit graph with $\gamma_0 = 0.440$, $\gamma_1 = 0.705$, and $\beta = 0.5$. The blue line shows $\langle H_p(t) \rangle$ according to the Schrödinger equation. The red line shows the Gaussian Ansatz, shown in Eq. 5.44, with τ marked by the dashed green line on the plot. Even though the system has been initialised in a Gibbs state, the Gaussian Ansatz still provides a good model. The dashed purple line in the figure shows τ^{num} from solving

$$\langle H_p(\tau^{\text{num}}) \rangle = \overline{\langle H_p \rangle} + \left(\langle H_p(0) \rangle - \overline{\langle H_p \rangle} \right) e^{-1}. \quad (5.47)$$

In Fig. 5.7a the value of τ^{num} is calculated numerically for various prob-

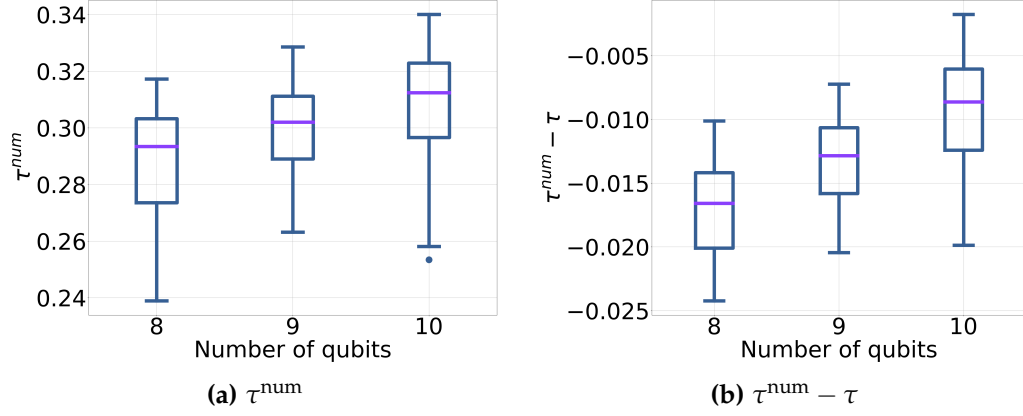


Figure 5.7: The thermalisation timescale for a closed-system CTQW on n qubits. For each instance, 100 examples are considered. The initial state is a Gibbs state with $\beta = 0.5$ and $\gamma_0 = (3/4)\sqrt{(n/\kappa_2)}$. For each instance $\gamma_1 = (6/5)\sqrt{(n/\kappa_2)}$.

lem instances and compared to Eq. 5.46 in Fig. 5.7b. For each problem instance $\gamma_0 = (3/4)\sqrt{(n/\kappa_2)}$, $\gamma_1 = (6/5)\sqrt{(n/\kappa_2)}$ and $\beta = 0.5$. The results are very similar to the case where the initial state was given by $|+\rangle$, shown in Fig. 5.4.

5.6 Conclusion

In this chapter, we have established that a reasonable approximation to the timescale associated with thermalisation of $\langle H_p \rangle$ in a closed quantum system is given by:

$$\frac{1}{\tau^2} = \frac{\text{Tr}([H_{QW}, [H_{QW}, H_p]] \rho_0)}{2(\langle H_p(0) \rangle - \overline{\langle H_p \rangle})}. \quad (5.48)$$

Importantly, this is an observable specific timescale. The physical intuition behind this timescale can be seen by identifying:

$$\text{Tr}([H_{QW}, [H_{QW}, H_p]] \rho_0) = \sum_{\omega_{m,n}} \omega_{m,n}^2 \nu_{m,n}, \quad (5.49)$$

as a measure of how spread the $\nu_{m,n}$ are in energy. The denominator can be identified with

$$\left(\langle H_p(0) \rangle - \overline{\langle H_p \rangle}\right) = \sum_{\omega_{m,n}} \nu_{m,n}, \quad (5.50)$$

how far the initial distribution is from equilibrium. On dimensional grounds alone it follows that

$$\tau \sim \sqrt{\frac{\sum \omega_{m,n} \nu_{m,n}}{\sum \omega_{m,n}^2 \nu_{m,n}}}. \quad (5.51)$$

That is to say τ is a rough measure of the square root of the height, to the width of the distribution of $\nu_{m,n}$ (see Fig. 5.2 for an example distribution). Therefore, we believe that this timescale will be a reasonable approximation in a wide range of circumstances and is easily adaptable to local other observables. To capture the dynamics of the CTQW a simple Ansatz has been chosen; there is scope for improving this Ansatz.

In this dissertation, the timescale associated with thermalisation is considered separately to the performance associated with the steady-state. It might be the case that sampling before the thermalisation occurs results in a better time-to-solution. From the discussion in this chapter, it is expected that the thermalisation timescale does not scale strongly with n . The time-to-solution is then mainly dependent on the ground-state probability, which we expect to be better once the system is thermalised (since this corresponds to a lower $\langle H_p \rangle$).

This chapter concludes the discussion of CTQWs, an example of a CTQO algorithm in a closed quantum system. As discussed in Chapter 2 there are other CTQO algorithms, which are time-dependent, and therefore do not correspond to a closed quantum system. The next chapter investigates these CTQO algorithms in an isolated quantum system, using CTQWs as a building block.

Statistical-mechanics applied to continuous-time quantum optimisation

6.1 Introduction

In Chapter 4 pure-state statistical physics was applied to CTQWs to make predictions. In this chapter, statistical physics is applied to a range of time-dependent quantum algorithms for optimisation that were described in Chapter 2. The algorithms are split into two broad classes: forward and cyclic approaches. Multi-stage quantum walks are an example of a forward quantum algorithm, where the system transitions from some initial Hamiltonian to a distinct final Hamiltonian. The step-like nature of MSQWs make them amenable for analysis by statistical physics, therefore MSQWs will be the focus of Sec. 6.2 as an example of a forward quantum algorithm. The continuum-limit of MSQWs is explored in Sec. 6.3. Cyclic approaches, such as Reverse Quantum Annealing (RQA), start and finish with the system experiencing the problem Hamiltonian. In Sec. 6.4 the limitations of cyclic approaches and computational mechanisms of cyclic approaches are explored. Central to the analysis in this section is Planck's principle, described in detail in Sec. 3.5.2. In short, an isolated quantum system can only increase in energy (i.e. no work can be extracted) as the result of a cyclic process.

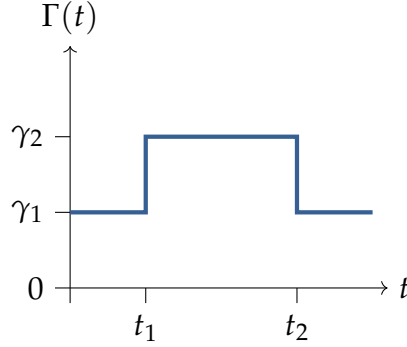


Figure 6.1: The cyclic schedule used to demonstrate the performance of MSQWs.

6.2 Multi-stage quantum walks

An MSQW is given by the following Hamiltonian:

$$H_{MS} = H_d + \Gamma(t)H_p, \quad (6.1)$$

where the schedule $\Gamma(t)$ is a monotonically increasing, piece-wise constant function. By applying Planck's principle (from Sec. 3.5.2), Sec. 6.2 shows how multi-stage quantum walks can be motivated. The thermal model introduced in Chapter 4 for CTQWs is then applied to the MSQW regime. Sec. 6.2.3 discusses what happens as more stages are added to an MSQW. Special attention is paid to passive states, which were introduced in Sec. 3.5.2. In Sec. 6.2.4 it is proven that passive states satisfy $\langle H_p(t) \rangle \leq \langle H_p(t = 0) \rangle$. Finally, in Sec. 6.2.5 a bath is introduced.

6.2.1 Motivating multi-stage quantum walks

This section demonstrates that $\langle H_p \rangle$ for an MSQW can only decrease as γ is increased, under Planck's principle. This motivates the choice of schedule $\Gamma(t)$ for MSQWs. To see this, consider the following cyclic

process with Hamiltonian H_{MS} and schedule:

$$\Gamma(t) = \begin{cases} \gamma_1 & \text{for } 0 \leq t \leq t_1 \\ \gamma_2 & \text{for } t_1 < t \leq t_2 \\ \gamma_1 & \text{for } t_2 < t \end{cases} \quad (6.2)$$

with $\gamma_2 - \gamma_1 = \Delta\gamma > 0$. The schedule is sketched out in Fig. 6.1. The energy at each stage is given by:

$$\langle H_{MS}(t) \rangle = \begin{cases} \langle H_{MS}(0) \rangle & \text{for } 0 \leq t \leq t_1 \\ \langle H_{MS}(0) \rangle + \Delta\gamma \langle H_p(t_1) \rangle & \text{for } t_1 < t \leq t_2 \\ \langle H_{MS}(0) \rangle + \Delta\gamma (\langle H_p(t_1) \rangle - \langle H_p(t_2) \rangle) & \text{for } t_2 < t. \end{cases}$$

The extractable work from the cyclic process, given by the initial energy minus the final energy, is:

$$W = \Delta\gamma (\langle H_p(t_2) \rangle - \langle H_p(t_1) \rangle). \quad (6.3)$$

If work cannot be extracted from the system, then $W \leq 0$ and

$$\langle H_p(t_2) \rangle \leq \langle H_p(t_1) \rangle. \quad (6.4)$$

Hence, if at each stage completing the cyclic process to return $\Gamma(t)$ to its value at the previous stage results in no extractable work, then $\langle H_p \rangle$ must be non-increasing between stages. Since at each stage the system approaches an equilibrium, characterised by Assumption 2, discussed in Sec. 3.3, this is expected to hold. During each stage, fluctuations will cause $\langle H_p \rangle$ to fluctuate in time around the stationary value (but not greater than the pre-quench value). If the initial stage corresponded to random guessing, as is the case in conventional MSQWs, this has as the corollary that one can always do better than random guessing. In summary, the expectation of H_p should be non-increasing with each stage.

Given significant dynamics, it is reasonable to expect $W < 0$ and $\langle H_p \rangle$ to decrease at each stage. But $\langle H_p \rangle$ is bounded, so cannot decrease forever. As $\Gamma(t)$ becomes very large, H_{MS} can be reasonably approximated as $\Gamma(t)H_p$, no longer driving transitions between eigenstates of H_p . Further to this, from Assumption 2 (Sec. 3.3), S_d increases with each stage. This means that $S_d = 0$ (corresponding to an eigenstate of H_p in the large $\Gamma(t)$ limit) becomes entropically forbidden if the system thermalises to any state with non-zero diagonal entropy.

From pure-state statistical physics, we have been able to reason about a typical system without resorting to numeric diagonalisation or recourse to adiabaticity. Assuming only that work cannot be extracted from a cyclic process, MSQWs have been motivated. As long as this always holds true, $\langle H_p \rangle$ can only decrease. At no point was the magnitude of the quenches specified, allowing for extension to QA-like schedules. This motivates monotonic schedules in QA.

6.2.2 A thermal model for multi-stage quantum walks

In Chapter 4 the tools for making predictions about CTQWs were developed. In this section, these tools are applied to the MSQW. Only the EMG distribution is used to model the DOS in this section, since this also captures the Gaussian model. For each stage, while γ is held constant, energy is conserved. Therefore, the temperature at each stage is fixed by:

$$\text{Tr} \left[H_{QW}^{(k)} \rho_\beta \left(H_{QW}^{(k)} \right) \right] = \langle \psi(t_{k-1}) | H_{QW}^{(k)} | \psi(t_{k-1}) \rangle, \quad (6.5)$$

where $H_{QW}^{(k)}$ is the Hamiltonian during the k^{th} stage of the quantum walk (i.e. with γ_k) and $|\psi(t)\rangle$ is the state vector associated with the MSQW. The term $\rho_\beta \left(H_{QW}^{(k)} \right)$ denotes a Gibbs state with inverse temperature β and Hamiltonian $H_{QW}^{(k)}$. Fig. 6.2 shows a five-stage MSQW for an instance of a 12-qubit binomial graph. Each stage increases γ by 0.5, starting with 0.5. Each stage has a duration of twenty units of

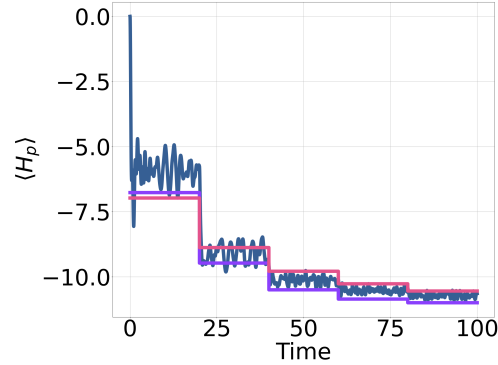


Figure 6.2: A five-stage CTQW on a 12-qubit binomial graph. The purple line is the numerical prediction, assuming that the MSQW is well modelled by a CTQW. The pink line shows the prediction from Eq. 4.26 and modelling the DOS as an EMG distribution.

time. With each successive stage, $\langle H_p \rangle$ (the solid blue line) improves and quickly tends to an approximate steady-state. Solving Eq. 6.5 to find the temperature and associated performance gives the purple line in the figure. For this instance, there is good qualitative and reasonable quantitative agreement between the thermal prediction and the Schrödinger equation. During the first stage of the MSQW, the thermal prediction is quite far from the prediction of the Schrödinger equation, suggesting that the state is far from thermal. This is to be expected, as γ is small and the driver Hamiltonian dominates, breaking the assumption that the DOS is continuous. Despite this, as γ is increased, the system thermalises. Perhaps most interestingly, as γ is further increased, the state remains thermal despite the problem Hamiltonian becoming more dominant. The inverse temperatures associated with this MSQW are $\beta = 1.21, 0.72, 0.53, 0.41, 0.32$. This corresponds to heating the system, suggesting that despite the final stage providing the best performance, it might be the distribution easiest to simulate classically [145]. This raises a question over where the advantage might be with this approach. It might be that the advantage can be found in the athermal behaviour not captured by approximating the pure state as a thermal state. Or that the MSQWs can prepare the thermal state faster than classical approaches by evolving through states that might not be classically simulable, re-

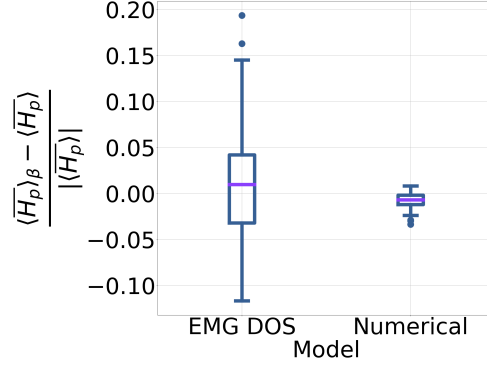


Figure 6.3: The difference between the stationary state calculated from the Schrödinger equation and a thermal state for five-stage CTQW on 100 12-qubit binomial graphs. “EMG DOS” refers to the prediction with the energy fixed by Eq. 6.6. The numerical approach gives the prediction from using the numerically determined energy, numerically fixing the temperature and calculating $\langle H_p \rangle$.

sulting in some less than exponential speed up. Given the thermal behaviour, it might be reasonable to assume that the techniques developed earlier in this dissertation for CTQWs can be applied to MSQWs.

To make analytic predictions, Eq. 6.5 needs to be approximated to find β . Since finding $|\psi(t)\rangle$ is likely to be numerically intractable for large systems, Eq. 6.5 can be approximated by:

$$\begin{aligned}
 & \langle \psi(t_{k-1}) | H_{QW}^{(k)} | \psi(t_{k-1}) \rangle \\
 &= \langle \psi(t_{k-1}) | H_d + \gamma_k H_p | \psi(t_{k-1}) \rangle \\
 &= \langle \psi(t_{k-1}) | H_d + \gamma_{k-1} H_p | \psi(t_{k-1}) \rangle \\
 & \quad + (\gamma_k - \gamma_{k-1}) \langle \psi(t_{k-1}) | H_p | \psi(t_{k-1}) \rangle \\
 &\approx \langle H_{QW}^{(k-1)} \rangle_{(k-1)} + (\gamma_k - \gamma_{k-1}) \overline{\langle H_p \rangle}_{(k-1)}, \tag{6.6}
 \end{aligned}$$

where $\langle \cdot \rangle_{k-1}$ denotes the expectation during the $(k-1)^{\text{th}}$ stage. Therefore, Eq. 6.5 can be estimated by recursively evaluating Eq. 6.6 using Eq. 4.26 to calculate $\overline{\langle H_p \rangle}_{(k-1)}$. In Fig. 6.2 the solid pink line shows the prediction, using Eq. 6.6, to good agreement.

In Fig. 6.3 the final value of $\overline{\langle H_p \rangle}$ for a hundred five-stage CTQW on 12-qubit binomial graphs is compared to the analytical prediction (i.e.

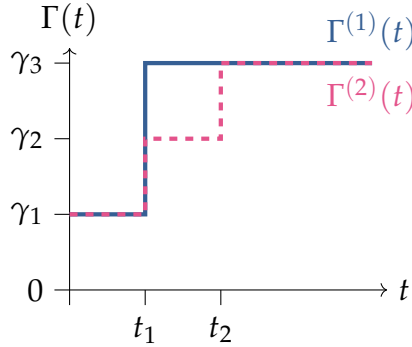


Figure 6.4: The monotonic schedules used to demonstrate that the use of more stages results in a lower final energy.

Eq. 6.6) and the numerical prediction assuming thermalisation. The schedule is the same as the schedule used in Fig. 6.2. Consider first the numerical prediction. Assuming thermalisation, the correct final steady state for all the problem instances is predicted within an error of 5%. In contrast, the EMG DOS model only achieves this error bound for approximately 50% of the instances. This approach suffers from cumulative errors in $\langle H_p \rangle$ at each stage of the MSQW (e.g. Eq. 6.6), which can result in large errors in the prediction.

This section has presented numerical evidence for MSQW exhibiting thermalisation. The EMG DOS model struggles to capture the performance for all instances, as well as the time-independent case. This is due to the difficulty in determining the energy of the system. To improve this approach, a better model for the DOS or prediction of $\langle H_p \rangle$ needs to be developed, especially for the small γ case.

6.2.3 How many stages?

Sec. 6.2.1 demonstrated that, under Planck's Principle (i.e. Assumption 4 detailed in Sec. 3.5.2) increasing from one value of γ to a larger value of γ decreases the value of $\langle H_p \rangle$. Sec. 6.2.2 demonstrated how the performance at each stage of the MSQW can be estimated. In this section, the potential benefit of adding an intermediate value of γ is discussed.

First, consider the schedule:

$$\Gamma^{(1)}(t) = \begin{cases} \gamma_1 & \text{for } 0 \leq t \leq t_1 \\ \gamma_3 & \text{for } t_1 < t \end{cases} \quad (6.7)$$

The associated energy for this schedule is:

$$\langle H_{MS}^{(1)}(t) \rangle = \begin{cases} \langle H_{MS}(0) \rangle & \text{for } 0 \leq t \leq t_1 \\ \langle H_{MS}(0) \rangle + (\gamma_3 - \gamma_1) \langle H_p(t_1) \rangle & \text{for } t_1 < t. \end{cases} \quad (6.8)$$

Consider a second schedule with an intermediate stage:

$$\Gamma^{(2)}(t) = \begin{cases} \gamma_1 & \text{for } 0 \leq t \leq t_1 \\ \gamma_2 & \text{for } t_1 < t \leq t_2 \\ \gamma_3 & \text{for } t_2 < t \end{cases} \quad (6.9)$$

where $\gamma_1 < \gamma_2 < \gamma_3$ and $t_1 < t_2$. Both schedules are shown in Fig. 6.4. The initial state for both schedules is assumed to be the same. A superscript (1) is used to distinguish quantities associated with the first schedule from quantities associated with the second schedule, superscripted by (2). If the quantity is the same for both schedules, the superscript is dropped. The second schedule has an associated energy:

$$\langle H_{MS}^{(2)}(t) \rangle = \langle H_{MS}(0) \rangle + \begin{cases} 0 & \text{for } 0 \leq t \leq t_1 \\ (\gamma_2 - \gamma_1) \langle H_p(t_1) \rangle & \text{for } t_1 \leq t \leq t_2 \\ (\gamma_3 - \gamma_2) \langle H_p^{(2)}(t_2) \rangle + (\gamma_2 - \gamma_1) \langle H_p(t_1) \rangle & \text{for } t_2 < t. \end{cases} \quad (6.10)$$

It follows that for $t > t_2$ the difference in energy between the two sched-

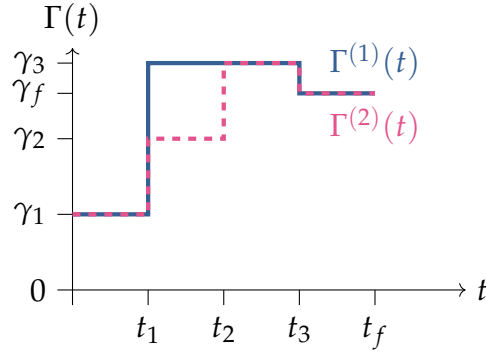


Figure 6.5: The schedules used to demonstrate that more stages result in a lower value of $\langle H_p \rangle$.

ules is:

$$\begin{aligned} \langle H_{MS}^{(2)}(t > t_2) \rangle - \langle H_{MS}^{(1)}(t > t_2) \rangle = \\ (\gamma_3 - \gamma_2) \left(\langle H_p^{(2)}(t_2) \rangle - \langle H_p(t_1) \rangle \right), \end{aligned} \quad (6.11)$$

since $\gamma_3 > \gamma_2$ by construction and $\langle H_p^{(2)}(t_2) \rangle \leq \langle H_p(t_1) \rangle$ (as argued in the previous section):

$$\langle H_{MS}^{(2)}(t > t_2) \rangle \leq \langle H_{MS}^{(1)}(t > t_2) \rangle. \quad (6.12)$$

Therefore, the energy of the system can only be decreased by the introduction of an intermediate stage.

Intuitively, a lower energy at the same value of Γ is likely to result in a smaller value of $\langle H_p \rangle$. The rest of this section attempts to substantiate this intuition. To do this, a criterion based on the energy of the schedules is set up, that if satisfied implies that a lower energy at the same Γ corresponds to a lower $\langle H_p \rangle$. Then from the density-of-states it is argued that this criterion is likely satisfied. For the schedules $\Gamma^{(1)}(t)$ and $\Gamma^{(2)}(t)$ consider an additional quench from γ_3 to γ_f at time t_3 , with $\gamma_f < \gamma_3$. The schedule is shown in Fig 6.5.

The change in energy for the modified $\Gamma^{(1)}$ schedule is

$$\begin{aligned} \langle (H_{MS}^{(1)}(t_f)) \rangle - \langle (H_{MS}^{(1)}(0)) \rangle = \\ (\gamma_3 - \gamma_1) \langle H_p(t_1) \rangle - (\gamma_3 - \gamma_f) \langle H_p^{(1)}(t_3) \rangle. \end{aligned} \quad (6.13)$$

The change in energy for the modified $\Gamma^{(2)}$ schedule is

$$\begin{aligned} \langle (H_{MS}^{(2)}(t_f)) \rangle - \langle (H_{MS}^{(2)}(0)) \rangle = (\gamma_3 - \gamma_2) \langle H_p^{(2)}(t_2) \rangle \\ + (\gamma_2 - \gamma_1) \langle H_p(t_1) \rangle - (\gamma_3 - \gamma_f) \langle H_p^{(2)}(t_3) \rangle \end{aligned} \quad (6.14)$$

The difference in the final energy between the two schedules is:

$$\begin{aligned} \langle H_{MS}^{(2)}(t_f) \rangle - \langle H_{MS}^{(1)}(t_f) \rangle = \\ (\gamma_3 - \gamma_2) \left(\langle H_p^{(2)}(t_2) \rangle - \langle H_p(t_1) \rangle \right) \\ - (\gamma_3 - \gamma_f) \left(\langle H_p^{(2)}(t_3) \rangle - \langle H_p^{(1)}(t_3) \rangle \right). \end{aligned} \quad (6.15)$$

Compare this with Eq. 6.11, the energy difference before the final quench. If the final quench has the effect of closing the energy gap between the two schedules, i.e.

$$\begin{aligned} \langle H_{MS}^{(2)}(t > t_f) \rangle - \langle H_{MS}^{(1)}(t > t_f) \rangle \\ - \left(\langle H_{MS}^{(2)}(t_f > t > t_2) \rangle - \langle H_{MS}^{(1)}(t_f > t > t_2) \rangle \right) \geq 0 \end{aligned} \quad (6.16)$$

then:

$$\langle H_p^{(2)}(t_3) \rangle \leq \langle H_p^{(1)}(t_3) \rangle. \quad (6.17)$$

This sets the criterion for determining if a lower energy at the same Γ corresponds to a lower $\langle H_p \rangle$. Now we will argue that this criterion is likely to be satisfied. From the Hamiltonian (Eq. 6.1) we might reasonably suspect that increasing $\Gamma(t)$ increases the spread of energy eigenstates, and conversely decreasing $\Gamma(t)$ decreases the spread. The vari-

ance in the energy of the eigenstates of H_{MS} is given by:

$$\sigma^2 = \frac{1}{\mathcal{D}} \left(\sum_k E_k^2 - \frac{1}{\mathcal{D}} \left(\sum_k E_k \right)^2 \right) \quad (6.18)$$

$$= \frac{1}{\mathcal{D}} \left(\text{Tr} (H_{QW}^2) - \frac{1}{\mathcal{D}} \text{Tr} (H_{QW})^2 \right), \quad (6.19)$$

where \mathcal{D} is the dimension of the underlying Hilbert-space. For brevity, let the normalised trace be denoted by $\text{Tr}' = 1/\mathcal{D} \text{Tr}$, and the scaled operators $\tilde{H}_p = H_p - \text{Tr}' H_p$ and $\tilde{H}_d = H_d - \text{Tr}' H_d$. Then,

$$\sigma^2 = \text{Tr}' \tilde{H}_p^2 \left(\Gamma + \frac{\text{Tr}' \tilde{H}_d \tilde{H}_p}{\text{Tr}' \tilde{H}_p^2} \right)^2 + \frac{\text{Tr}' \tilde{H}_d^2 \text{Tr}' \tilde{H}_p^2 - (\text{Tr}' \tilde{H}_d \text{Tr}' \tilde{H}_p)^2}{\text{Tr}' \tilde{H}_p^2}. \quad (6.20)$$

Therefore, provided

$$\Gamma > -\frac{\text{Tr}' \tilde{H}_d \tilde{H}_p}{\text{Tr}' \tilde{H}_p^2} \quad (6.21)$$

holds true, decreasing Γ decreases σ^2 . For many typical cases, such as H_p consisting of Pauli Z terms and H_d consisting of Pauli X terms, $\text{Tr}' \tilde{H}_d \tilde{H}_p = 0$. In this case, Eq. 6.21 reduces to $\Gamma > 0$.

Given that γ_f can be arbitrarily close to γ_3 , a very small decrease in Γ will bring the energy eigenstates closer together without significantly affecting the populations of each state. This will likely decrease the energy gap between the two schedules, implying Eq. 6.17.

This presents a trade-off between anneal-time and performance. More stages will result in a longer anneal time, given the requirement to thermalise at each stage, but an improved performance as demonstrated by a lower value of $\overline{\langle H_p \rangle}$ with each stage.

6.2.4 Consequences of passivity

In the previous sections, Planck's principle was used as a physically reasonable assumption to discuss CTQO, specifically MSQWs. As men-

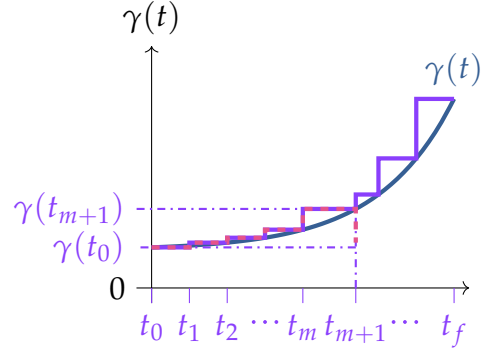


Figure 6.6: The solid blue line sketches a possible $\gamma(t)$, which in this case is monotonically increasing. The purple line shows the piece-wise constant function used to approximate $\gamma(t)$. The dashed pink line shows the cyclic process used to demonstrate $\langle H_p(t) \rangle \leq \langle H_p(0) \rangle$.

tioned in Sec. 3.5.2, Planck's principle is provably true for arbitrary cyclic unitaries and passive states. The aim of this section is to make clear a provable statement that follows from passivity (introduced in Sec. 3.5.2). This section allows for the initial state to be a mixed state, but the system is otherwise isolated. The Hamiltonian under consideration is

$$H(t) = H_d + \gamma(t)H_p, \quad (6.22)$$

with, t varying from $t = t_0$ to $t = t_f$. The initial state is any passive state with Hamiltonian $H(t_0)$. The schedule $\gamma(t)$ is approximated with a step-wise constant function, akin to an MSQW, shown in Fig. 6.6. The change in energy as a result of the piece-wise constant function is

$$\langle H(t_f) \rangle = \sum_{k=0}^{f-1} [\gamma(t_{k+1}) - \gamma(t_k)] \langle H_p(t_k) \rangle. \quad (6.23)$$

At the end of the anneal a fictitious quench is introduced to return the Hamiltonian to $H(t_0)$. The extractable work from this process is

$$W = - \sum_{k=0}^{f-1} [\gamma(t_{k+1}) - \gamma(t_k)] \langle H_p(t_k) \rangle + [\gamma(t_f) - \gamma(t_0)] \langle H_p(t_f) \rangle. \quad (6.24)$$

Since the initial state is passive, i.e. $W \leq 0$, it follows that

$$\langle H_p(t_f) \rangle \leq \sum_{k=0}^{f-1} \left[\frac{\gamma(t_{k+1}) - \gamma(t_k)}{\gamma(t_f) - \gamma(t_0)} \right] \langle H_p(t_k) \rangle. \quad (6.25)$$

That is to say, $\langle H_p(t_f) \rangle$ is bounded by a weighted average over the previous stages. Consider now a monotonically increasing schedule for $\gamma(t)$, such that $\gamma_{k+1} > \gamma_k$. For $k = 0$, and introducing a fictitious quench at t_1 to create a cyclic process, it follows that:

$$W_{10} = [\gamma(t_1) - \gamma(t_0)] (\langle H_p(t_1) \rangle - \langle H_p(t_0) \rangle) \leq 0, \quad (6.26)$$

hence

$$\langle H_p(t_1) \rangle \leq \langle H_p(t_0) \rangle. \quad (6.27)$$

Assuming that

$$\langle H_p(t_k) \rangle \leq \langle H_p(t_0) \rangle \quad (6.28)$$

is true for all $k \leq m$, it remains to show that it is true for $k = m + 1$. Again a fictitious quench is introduced at t_{m+1} to create a cyclic process:

$$\begin{aligned} W_{m+1,0} = & - \sum_{k=0}^m [\gamma(t_{k+1}) - \gamma(t_k)] \langle H_p(t_k) \rangle \\ & + [\gamma(t_{m+1}) - \gamma(t_0)] \langle H_p(t_{m+1}) \rangle. \end{aligned} \quad (6.29)$$

Again, due to passivity, $W_{m+1,0} \leq 0$. Therefore:

$$\langle H_p(t_{m+1}) \rangle \leq \sum_{k=0}^m \left[\frac{\gamma(t_{k+1}) - \gamma(t_k)}{\gamma(t_{m+1}) - \gamma(t_0)} \right] \langle H_p(t_k) \rangle, \quad (6.30)$$

from the assumption stated in Eq. 6.28

$$\langle H_p(t_{m+1}) \rangle \leq \underbrace{\langle H_p(t_0) \rangle \sum_{k=0}^m \left[\frac{\gamma(t_{k+1}) - \gamma(t_k)}{\gamma(t_{m+1}) - \gamma(t_0)} \right]}_{=1}. \quad (6.31)$$

By induction

$$\langle H_p(t_k) \rangle \leq \langle H_p(t_0) \rangle \quad (6.32)$$

for all $k \geq 0$ given a monotonically increasing $\gamma(t_k)$. Making the time-step arbitrarily small and returning to the continuous limit gives:

$$\langle H_p(t) \rangle \leq \langle H_p(t_0) \rangle \quad (6.33)$$

for all $t \geq t_0$ given a monotonically increasing $\gamma(t)$. This has previously been shown for the case where the initial state is the ground state of H_d [67]. Here it has been extended to any passive state, including ground states and Gibbs states with Hamiltonian $H_d + \gamma(t_0)H_p$ that have a non-trivial value of $\langle H_p(t_0) \rangle$. This suggests that passive states might have a significant role to play in QA. Further to assurances considered so far, given any initial mixed state, the final state that minimises H_p varying over all possible unitaries must be passive with respect to H_p . If the final state is not passive, there exists a unitary that lowers the energy.

6.2.5 Introduction of a bath

To explore some consequences of open quantum system effects, a time-independent bath is added to the Hamiltonian with time-independent couplings to the quantum system. The resulting Hamiltonian is given by:

$$H_{\text{total}} = H_S(t) + H_{SB} + H_B, \quad (6.34)$$

where H_B acts on the bath, $H_S(t)$ on the system and H_{SB} is the system-bath interaction. For simplicity, all the operators are taken to be trace-class and the joint system to be isolated.

Consider the case where $H_S(t)$ is time-independent and given by $H_S(t) = H_d + \gamma H_p$. After some time, under the ETH (Assumption 3), the total system will be locally indistinguishable from a Gibbs state, with Hamiltonian H_{total} . The system's state, given by tracing out the bath, need not be a Gibbs state at all.

The temperature is fixed by the initial conditions:

$$\langle H_{\text{total}} \rangle = -\frac{\partial \log \mathcal{Z}_{\text{total}}}{\partial \beta}, \quad (6.35)$$

where

$$\mathcal{Z}_{\text{total}} = \text{Tr} \left(e^{-\beta H_{\text{total}}} \right). \quad (6.36)$$

As the system becomes small compared to the bath, the system can be neglected in Eq. 6.35. At this point β becomes a property of the bath and independent of γ .

The expectation value of $\langle H_p \rangle$ is given by

$$\langle H_{\text{total}} \rangle = -\frac{1}{\beta} \frac{\partial \log \mathcal{Z}_{\text{total}}}{\partial \gamma}. \quad (6.37)$$

On the assumption β is fixed such that it is independent of γ , under the Peierls-Bogoliubov inequality the free-energy is concave [155], i.e.,

$$-\frac{1}{\beta} \frac{\partial^2 \ln \mathcal{Z}_{\text{total}}}{\partial \gamma^2} \leq 0. \quad (6.38)$$

It follows that:

$$\frac{\partial \langle H_p \rangle}{\partial \gamma} \leq 0, \quad (6.39)$$

so $\langle H_p \rangle$ is a monotonically decreasing function of γ for a fixed temperature Gibbs state. It follows that the optimal choice of γ is as large as possible. Beyond a certain point, the assumption that β is independent of γ breaks down. Note that this argument only made use of the fact that the joint system is locally indistinguishable from a Gibbs state, making no assumptions about the form of the terms beyond being Hermitian. In summary, optimising Gibbs states with fixed temperature is straightforward and is problem-instance independent. This simplicity is in stark contrast to the bath-free case [22]. Utilising a bath to solve optimisation problems has previously been explored by Imperato *et al.* [156]. They used auxiliary qubits as a bath to tune the effective temperature of the

joint-system. The explicit measurement of temperature is, however, not touched upon.

Further to this, once the joint system has approached thermal equilibrium, if the system is then quenched such that $\gamma(t)$ is increased at $t = 0$, it follows from the passivity of the Gibbs state that $\langle H_p(t) \rangle \leq \langle H_p(0) \rangle$. This suggests some inbuilt error resistance to CTQO. The inbuilt resistance to errors in QA has been explored in [84, 85, 86].

6.3 Pure State Thermal Quantum Annealing

In this section, the continuum limit of an MSQW is considered, such that the rate at which $\Gamma(t)$ changes is sufficiently slow compared to the dephasing timescale of the system. By applying the ETH (Assumption 3), the system is then modelled as a Gibbs state at all times. The coefficient in front of the driver Hamiltonian is allowed to change too, under the same assumptions. The Hamiltonian is given by:

$$H_{TQA} = A(t)H_d + B(t)H_p, \quad (6.40)$$

where $A(t)$ and $B(t)$ are changed slowly compared to the dephasing timescale. Intuition into this timescale can be found in [109, 110, 111], where it is argued that this timescale is unlikely to be exponential. Detailed numerical studies focusing on the systems considered in this dissertation, including the role of the initial state, have not been carried out yet. However, provided that the state (including the initial state) consists of an exponential number of energy eigenstates we expect that the dephasing timescale will be less than exponential, independent of the state. This is consistent with the result for CTQWs discussed in Chapter 5. That is not to say that the dephasing time is independent of the state, as we saw in Chapter 5, the dephasing timescale depended on the difference between the initial value of the observable (corresponding to the initial state) and the infinite time-averaged value. In contrast, adiabatic quantum optimisation has overlap with a single energy eigenstate and

an exponentially scaling runtime.

We refer to this limit as pure state thermal quantum annealing (PSTQA). Since the system is isolated:

$$\frac{d\langle H_{TQA}(t) \rangle}{dt} = \left\langle \frac{\partial H_{TQA}}{\partial t} \right\rangle \quad (6.41)$$

$$= \dot{A}(t)\langle H_d \rangle + \dot{B}(t)\langle H_p \rangle, \quad (6.42)$$

where the dot denotes a time derivative. Denoting the time-dependent partition function of the system as $\mathcal{Z}(t)$:

$$\langle H_{TQA}(t) \rangle = - \frac{\partial \ln \mathcal{Z}(t)}{\partial \beta} \quad (6.43)$$

$$\langle H_d(t) \rangle = - \frac{1}{\beta(t)} \frac{\partial \ln \mathcal{Z}(t)}{\partial A} \quad (6.44)$$

$$\langle H_p(t) \rangle = - \frac{1}{\beta(t)} \frac{\partial \ln \mathcal{Z}(t)}{\partial B}, \quad (6.45)$$

where β is the inverse temperature. The rest of this section investigates the consequences of these equations.

6.3.1 Pure state thermal quantum annealing is path-independent

Returning to Eq. 6.42, the equation can be rewritten into normalised time with the substitution $s = t/t_f$, where t_f is the anneal time,

$$\frac{d\langle H_{TQA}(s) \rangle}{ds} = \langle H_d \rangle \frac{dA(s)}{ds} + \langle H_p \rangle \frac{dB(s)}{ds}. \quad (6.46)$$

The above equation depends only on the normalised time and the expectation values depend only on A, B and β and not their derivatives. This implies $\langle H_p \rangle$ in the thermal model is independent of scaling by t_f , although obviously too small a t_f will cause the thermal predictions to

break down. Further rewriting Eq. 6.46 gives:

$$\langle H_{TQA}(1) \rangle = \int_{A(0), B(0)}^{A(1), B(1)} \begin{pmatrix} \langle H_d \rangle \\ \langle H_p \rangle \end{pmatrix} \cdot \begin{pmatrix} dA \\ dB \end{pmatrix}, \quad (6.47)$$

hence Eq. 6.46 can be viewed as a path integral under the “force”

$$\vec{D} = - \begin{pmatrix} \langle H_d \rangle \\ \langle H_p \rangle \end{pmatrix} = -\vec{\nabla} F, \quad (6.48)$$

where F is the Helmholtz free energy, $F = -\ln(\mathcal{Z})/\beta$ and $\vec{\nabla} = (\partial_A, \partial_B)^T$. Since each of $\langle H_d \rangle$ and $\langle H_p \rangle$ depend only on A and B and not their derivatives and the force can be written as the gradient of a scalar field, $\langle H_{TQA} \rangle$ is path independent. It follows that $\langle H_p \rangle$ will also be path-independent.

6.3.2 Diagonal entropy is conserved

The diagonal entropy, assuming a Gibbs state, is given by

$$S_d(\beta, E, A, B) = \ln \mathcal{Z}(\beta, A, B) + \beta E, \quad (6.49)$$

where $E = \langle H_{TQA}(t) \rangle$. Taking the time derivative of S_d gives

$$\frac{dS_d}{dt} = \frac{\partial \ln \mathcal{Z}}{\partial A} \frac{dA}{dt} + \frac{\partial \ln \mathcal{Z}}{\partial B} \frac{dB}{dt} + \left(\frac{\partial \ln \mathcal{Z}}{\partial \beta} + E \right) \frac{d\beta}{dt} + \beta \frac{dE}{dt}, \quad (6.50)$$

substituting in Eq. 6.43 gives:

$$\frac{dS_d}{dt} = \frac{\partial \ln \mathcal{Z}}{\partial A} \frac{dA}{dt} + \frac{\partial \ln \mathcal{Z}}{\partial B} \frac{dB}{dt} + \beta \frac{dE}{dt}, \quad (6.51)$$

Using Eqs. 6.44, 6.45 and Eq. 6.42:

$$\frac{dS_d}{dt} = 0. \quad (6.52)$$

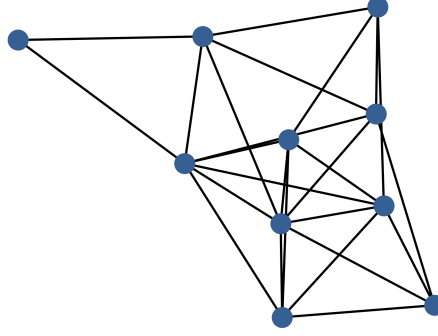


Figure 6.7: The 10-qubit Max-cut example.

In conclusion, PSTQA is the adiabatic limit of an MSQW, with Assumption 3. The standard quantum adiabatic theorem would imply that the relevant timescale would be associated with the minimum spectral gap, hence a timescale that grows exponentially with the number of qubits. But to reach PSTQA the relevant timescale invoked was the dephasing timescale. To make Assumption 3, it was assumed that non-extensive corrections to the observables could be neglected. Since S_d is an extensive quantity, it is perhaps more reasonable to say that PSTQA is adiabatic up to non-extensive corrections.

6.3.3 Numerical evidence of PSTQA

To illustrate PSTQA this section numerically examines a specific example. The example is a single instance of Max-cut on a randomly generated graph with 10 qubits, with H_d being the transverse-field Hamiltonian. Each edge in the Max-cut graph has been selected with probability $2/3$. The Max-cut graph is shown in Fig. 6.7. The schedule is shown in Fig. 6.8, where $B(0) \neq 0$ and $A(t_f) \neq 0$ breaking integrability and the conventional assumption in AQO.

Fig. 6.9 shows $\langle H_p(t) \rangle$ for the Max-cut example. The blue line shows the Schrödinger equation. The pink line shows the prediction from directly solving the PSTQA equations (i.e. Eqs. 6.42-6.45). There is very good agreement throughout the evolution. Fig. 6.10 shows the diago-

6. STATISTICAL-MECHANICS APPLIED TO CONTINUOUS-TIME QUANTUM OPTIMISATION

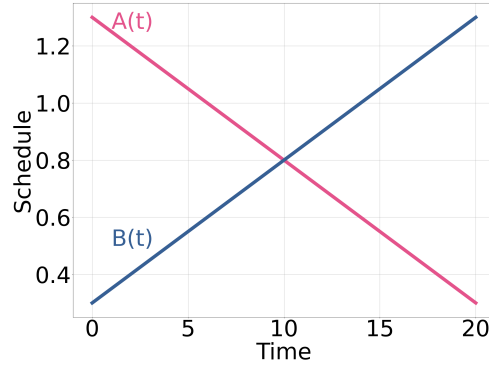


Figure 6.8: The schedule used for the 10-qubit Max-cut instance. The pink line shows $A(t)$ and the blue line shows $B(t)$.

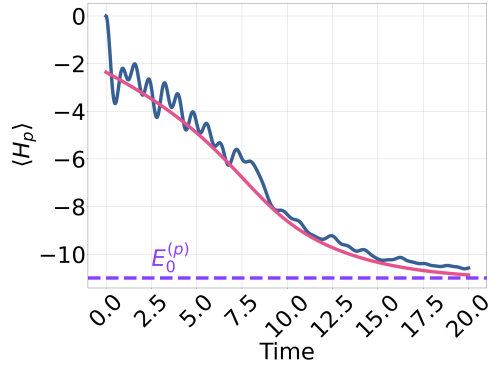


Figure 6.9: A comparison of $\langle H_p(t) \rangle$ calculated from the Schrödinger equation (blue) and the PSTQA equations (pink) for the 10 qubit example. The dashed purple line shows the ground-state energy, $E_0^{(p)}$, of the problem Hamiltonian.

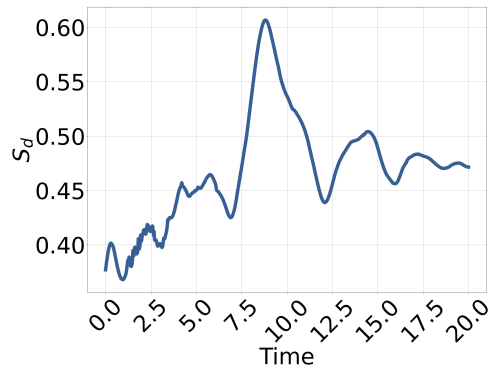


Figure 6.10: The diagonal entropy calculated from the Schrödinger equation for the 10 qubit Max-cut example.

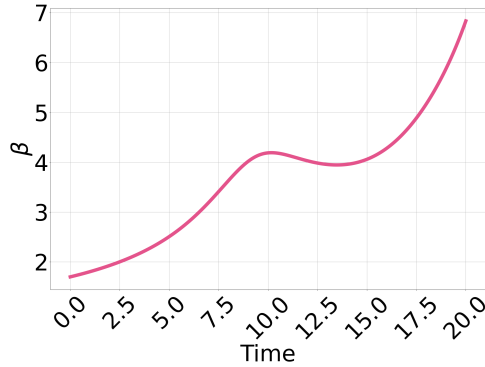


Figure 6.11: The inverse temperature calculated from the PSTQA equations for the 10 qubit Max-cut example.

nal entropy for the same evolution, according to the Schrödinger evolution. The change is relatively small, however non-zero. This is evidence of diabatic transitions between energy eigenstates. Indeed, at the end of the evolution, there is a net increase in diagonal entropy. The inverse temperature calculated from Eq. 6.43 is shown in Fig. 6.11. Generally, the temperature is seen to be decreasing, except at around $t \approx 11$ where the system heats. The peak in diagonal entropy (according to the Schrödinger equation) and inverse temperature (according to the PSTQA equations) appears to roughly coincide. These are not directly linked, recall that the PSTQA equations predict no change in diagonal entropy. However, halfway through the anneal the variance in energy eigenstates is minimised, providing a possible explanation of why numerous transitions occur resulting in rapid change in the diagonal entropy. Given an approximate Gaussian density of states for the evolution, see the next chapter and Appendix A.2) we expect that a maximum in the inverse temperature occurs when the variance is minimised. In Appendix A.1 more numerical evidence of PSTQA is provided, including a second combinatorial optimisation problem to show that this phenomenon is not restricted to Max-cut.

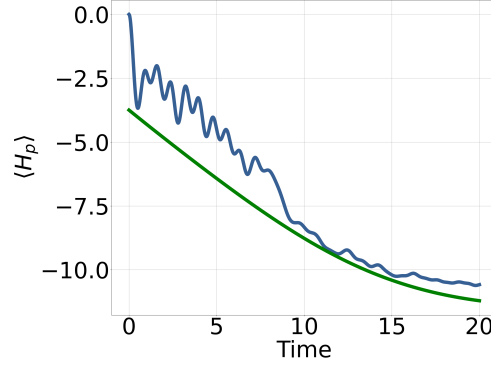


Figure 6.12: PSTQA for the 10-qubit Max-cut graph shown in Fig. 6.7. The blue line shows the Schrödinger evolution. The green line shows the prediction from the PSTQA equations when an exponentially modified Gaussian ansatz is made for the density-of-states.

6.3.4 Ansatz based approach

Directly numerically solving the PSTQA equations (i.e. Eqs. 6.42-6.45) is difficult, requiring repeated matrix exponentiation to find β . However, since the PSTQA equations depend only on the partition function, the equations can be tackled by making a good choice of ansatz for \mathcal{Z} . Ansätze for the partition function based on models of the density-of-states for Max-cut were explored in Chapter 4 for the time-independent setting. Fig. 6.12 shows the EMG ansatz for the density-of-states, for the time-dependent setting. The green line in Fig. 6.12 shows the ansatz-based approach, and the pink line again shows the Schrödinger evolution of $\langle H_p(t) \rangle$ for the 10-qubit instance considered in the previous section. Although the agreement is not as good as exactly solving the PSTQA equations, this approach removes the need for matrix exponentiation and still provides a good approximation. The details of the calculation can be found in Appendix A.2, along with further numerical examples.

6.4 Cyclic approaches

Cyclic processes present a natural formulation for incorporating prior classical information about the optimisation problem into the quantum algorithm. Instead of starting in a pure state, the system is initialised in a classical ensemble, diagonal in the computational basis. This use of prior information is sometimes known as warm-starting. Planck's principle is naturally formulated in terms of a cyclic process. Here, we investigate what this means for CTQO. Sec. 6.4.1 considers warm-started CTQWs. Sec. 6.4.2 considers reverse quantum annealing (RQA) more broadly. In Sec. 6.4.3 considers biased quantum annealing (BQA), focusing on the computational mechanism.

6.4.1 Warm starting continuous-time quantum walks does not work (on average)

From previous runs, or from some prior knowledge of the problem, a reasonable non-optimal solution to the combinatorial optimisation problem might be known. The case of a CTQW, where instead of starting in an eigenstate of H_d it is initialised in an eigenstate of H_p , is studied in this section. In order to apply Assumption 4, it is required that the initial state $|z^*\rangle$ is better than average (see Sec. 3.5.2 for further discussion):

$$\langle z^* | H_p | z^* \rangle < \frac{1}{2^n} \text{Tr } H_p. \quad (6.53)$$

Consider the Hamiltonian:

$$H_{WS} = G(t)H_d + H_p, \quad (6.54)$$

where

$$G(t) = \begin{cases} 0 & \text{for } t \leq 0 \\ g & \text{for } 0 < t \leq t_1 \\ 0 & \text{for } t_1 \leq t. \end{cases} \quad (6.55)$$

Note that the time-dependent schedule is now appended to H_d , not H_p . This cyclic process is indistinguishable from a warm-started CTQW period with a duration of t_1 . Again, Planck's Principle (Assumption 4) is applied. The extractable work is:

$$W = g (\langle H_d(t_1) \rangle - \langle H_d(0) \rangle) \leq 0. \quad (6.56)$$

We conclude that if $g > 0$, $\langle H_d(t_1) \rangle \leq \langle H_d(0) \rangle$. Else if $g < 0$, $\langle H_d(t_1) \rangle \geq \langle H_d(0) \rangle$. By conservation of energy

$$\langle H_p(0) \rangle + g \langle H_d(0) \rangle = \langle H_p(t_1) \rangle + g \langle H_d(t_1) \rangle. \quad (6.57)$$

Combining the above statements:

$$\langle H_p(0) \rangle - \langle H_p(t_1) \rangle = g (\langle H_d(t_1) \rangle - \langle H_d(0) \rangle) \leq 0. \quad (6.58)$$

Hence,

$$\langle H_p(0) \rangle \leq \langle H_p(t_1) \rangle, \quad (6.59)$$

under Assumption 4. Since $\langle H_p(0) \rangle$ corresponds to the initial guess, the CTQW gives on average worse-quality solutions than the initial guess. Therefore, CTQWs cannot be on average warm-started by simply driving from a low energy eigenstate of H_p . Numerical evidence of the performance of warm-started CTQWs can be found in Appendix A.3.

6.4.2 Reverse quantum annealing

In this section, reverse quantum annealing (RQA) is understood to be a cyclic protocol that maps a classical probability density function of computational basis states to another probability density function. A warm-started CTQW can be considered an example of RQA. Under Assumption 4, the total energy of the system has to rise as a result of a cyclic process. This rules out cooling H_p cyclically to find good quality

solutions given a Hamiltonian of the form

$$H_{cyc}(t) = H_p + G(t)H_d, \quad (6.60)$$

where $0 \leq t \leq t_{cyc}$ and $G(t)$ corresponds to a cyclic process (i.e. $G(t = 0) = G(t = t_{cyc}) = 0$). This is at odds with approaches such as RQA. Further to this, the following is provably true:

1. The entropy of the resulting distribution is greater than or equal to the entropy of the original distribution. The cyclic process causes the distribution to become more spread (see Chapter 3).
2. If the initial state is passive, for example a thermal state of H_p , then the mean of the resulting distribution can only increase. That is to say, if the initial state is passive, then there exists no unitary cyclic process that will reduce $\langle H_p \rangle$ averaged over the distribution.

Beyond $\langle H_p \rangle$, it is also possible to draw some conclusions about the ground-state probability for RQA. The initial state for RQA is the ensemble

$$\rho_0 = \sum_s p(s) |s\rangle \langle s|, \quad (6.61)$$

where $|s\rangle$ is an eigenstate of H_p with eigenvalue s . Denoting the unitary associated with one RQA cycle to be U_{cyc} , then the transition probability between states $|s\rangle$ and $|j\rangle$ is given by

$$p(j|s) = |\langle j| U_{cyc} |s\rangle|^2. \quad (6.62)$$

Note that the elements $p(j|s)$ constitute a doubly-stochastic matrix, \mathbf{P} [1]. From Birkhoff's theorem [1, 157], \mathbf{P} can be written as a convex sum of permutation matrices Π_α , such that the resulting state ρ_1 can be written as:

$$\rho_1 = \mathbf{P}\rho_0 \quad (6.63)$$

$$\rho_1 = \sum_\alpha q_\alpha \Pi_\alpha \rho_0 \quad (6.64)$$

where $0 \leq q_\alpha \leq 1$ and $\sum_\alpha q_\alpha = 1$. Here ρ_i , $i = 0, 1$ can be viewed as the diagonal part of the density operator in the computational basis. Interpreting q_α as a probability, RQA can then be viewed as applying a permutation Π_α with probability q_α . The distribution of q_α is determined by the schedule of the RQA protocol. This is likely to consist of some local and global moves, with locality being driver dependent.

The doubly-stochastic evolution places restrictions on the resulting distribution ρ_1 . As mentioned previously, if ρ_0 is passive, then \mathbf{P} can only reduce the average quality of solution. Further to this, looking at the change in ground-state probability $\Delta p(0)$ with initial ground-state probability $p(0)$:

$$\Delta p(0) = \sum_m p(0|m)p(m) - p(m|0)p(0) \quad (6.65)$$

$$\Delta p(0) = \sum_m p(0|m) (p(m) - p(0)) \quad (6.66)$$

$$(6.67)$$

where $p(0|m) = p(m|0)$ has been used. Bounding $\Delta p(0)$ gives:

$$\min_m (p(m) - p(0)) \sum_m p(0|m) \leq \Delta p(0) \leq \max_m (p(m) - p(0)) \sum_m p(0|m). \quad (6.68)$$

Since \mathbf{P} is doubly-stochastic,

$$\min_m (p(m) - p(0)) \leq \Delta p(0) \leq \max_m (p(m) - p(0)). \quad (6.69)$$

This means if the ground state is already the most-likely state (even if it is exponentially small) \mathbf{P} can only reduce the probability of finding the ground state. If the ground state is the least likely state, then RQA can only improve the ground-state probability. More generally, the ground-state probability is bounded by $\max_m p(m)$, which if exponentially small restricts finding the ground state to being exponentially unlikely. If $p(m)$ is sampled from a system with an extensive amount of entropy, then this is likely the case.

After each stage, there is some selection criterion to determine if a string is kept or if the RQA cycle is repeated with the initial string. The selection criterion could be the Metropolis-Hastings update rule [158] or a more complicated update [70, 71]. At its core, the quantum part of RQA is a method of generating q_α and implementing the corresponding permutation. It remains to be shown that this is better than classically generating a distribution of q_α and corresponding permutation, say for example with cluster Monte Carlo [159].

Note that an average increase in $\langle H_p \rangle$ does not rule out finding states with a lower value of $\langle H_p \rangle$ - this is simply not the average case. Since in RQA the system starts in an energy eigenstate, the uncertainty can only increase as a result of the cyclic process. Provided the shift in average energy is less than the increased uncertainty in energy, RQA has a chance of doing better than its initial guess.

6.4.3 Biased quantum annealing

In the previous section, it was shown that RQA is restricted in terms of which states it can reach. Since the energy of the system tends to increase a third term needs to be added to invalidate Planck's principle (this is discussed in detail later in this chapter),

$$H_{cyc}(t) = H_p + H_b + G(t)H_d. \quad (6.70)$$

Provided that the increase in energy associated with the bias Hamiltonian (H_b) is greater than the energy increase as a result of the cyclic process, the energy associated with H_p will decrease. To remove ambiguity, we refer to RQA with a bias as biased quantum annealing (BQA).

The following protocol is inspired by works such as [73] and [74]. In these works, the many-body localised phase transition in spin glasses is used to cyclically cool the system. We argue that this approach is likely to hold for a broad range of settings. Initialising in an eigenstate of H_p , denoted by $|z^*\rangle$ and $H_b = -\alpha |z^*\rangle \langle z^*|$ with $\alpha > 0$, lowers the energy of

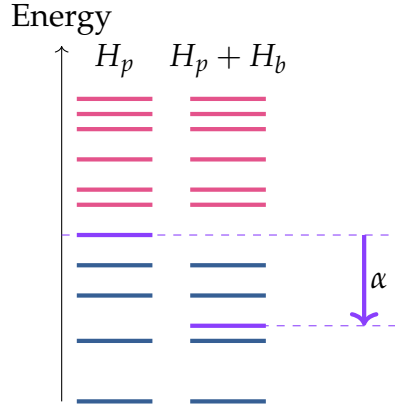


Figure 6.13: The effect on the problem Hamiltonian of introducing H_b . The initial state is highlighted in purple. States lower in energy are coloured blue, while states higher in energy are coloured pink. As long as $|\Delta W| < \alpha$, the expectation value of $\langle H_p \rangle$ will decrease.

the initial state, while all other eigenstates of H_p remain unchanged. As long as the energy increase ΔW from the cyclic process is smaller than the energy shift from the bias, $\langle H_p \rangle$ will be reduced. This is sketched out in Fig. 6.13. Note that for this choice of biasing Planck's principle no longer holds as a map between classical probability distributions, this is discussed in Sec. 6.4.3. However, we will assume that the net result of a typical cyclic process with sufficient dynamics is to move towards the bulk of the spectrum, increasing the energy. For the process to work, non-trivial dynamics need to take place. Given some starting state, the process could work as follows:

1. Run the cyclic process
2. If the resulting string corresponds to a better solution, update the initial state to be this state.
3. Otherwise, keep the initial starting state but increase α .

In practice, achieving a bias like $H_b = -\alpha |z^*\rangle \langle z^*|$ will be infeasible. Such a bias would involve all-to-all interactions. A much more feasible driver is

$$H_b^{(l)} = -\alpha \sum_{i=1}^n (-1)^{z^*[i]} Z_i, \quad (6.71)$$

where $z^*[i]$ is the i^{th} bit in the n -bit string z^* , and Z_i is the Pauli Z matrix acting on the i^{th} qubit. This Hamiltonian has the ground state $|z^*\rangle$ with energy $-n\alpha$, and consists of one-local terms. However, this Hamiltonian will alter the energy of all the problem states. In the worst case, the energy gap opened up by $H_b^{(l)}$ becomes primarily occupied by high energy states of H_p and no cooling of $\langle H_p \rangle$ occurs.

If however, assuming that the ordering of energy eigenstates is uncorrelated between $H_b^{(l)}$ and H_p , then the average energy shift from $H_b^{(l)}$ is 0 with standard deviation $\alpha\sqrt{n}$. So the typical energy shift of the problem eigenstates is \sqrt{n} times smaller than the shift on the initial state. Therefore, for large n in the typical case, we would expect this local bias to mimic $-|z^*\rangle\langle z^*|$. For this approach, if α is too big $H_b^{(l)}$ will dominate and the approach will not involve any information from H_p . The performance of this approach is numerically explored in Appendix A.4.

Does BQA violate the second law of thermodynamics?

We have discussed that if the system is likely to increase in energy, then the introduction of a bias can help improve $\langle H_p \rangle$. This section outlines how BQA can appear to violate the second-law of thermodynamics while elucidating the resources used in BQA.

Let's consider a cyclic process where the initial state

$$\rho_0 = \sum_z p_z |z\rangle\langle z| \quad (6.72)$$

is diagonal in the computational basis. For each run, the state $|z\rangle$ is fed into the quantum annealer with probability p_z . The initial Hamiltonian is H_p . The system interpolates between H_p and $H_b^{(l)}$ such that the system reaches the ground-state of $H_b^{(l)}$. The system is then adiabatically evolved to the ground-state of H_d (which by assumption commutes with neither H_p nor $H_b^{(l)}$). From here, the system is adiabatically evolved to the ground-state of H_p . The result of this cyclic process is to take ρ_0 and map it to the ground-state of the problem Hamiltonian. Hence, a cyclic

6. STATISTICAL-MECHANICS APPLIED TO CONTINUOUS-TIME QUANTUM OPTIMISATION

process that leads to extractable work and a decrease in entropy has been constructed, violating the second law of thermodynamics. This is true even for passive states. This process is typically referred to as adiabatic reverse quantum annealing [160, 161].

The apparent violation arises because the evolution is not unitary (hence not isolated), evidenced by the change in von Neumann entropy. The choice of cyclic process is predicated on the initial state loaded into the quantum annealer. To make the system isolated, consider the introduction of a second register of qubits. The first register Q contains the qubits used in the BQA process and contains n qubits. The second register C contains the same number of bits as Q , and each qubit in the second register is prepared in $|0\rangle$. The joint initial state is now:

$$\rho_0^{(Q/C)} = \sum_z p_z |z\rangle \langle z|_Q \bigotimes_{i=0}^{n-1} |0\rangle \langle 0|_{i,C}. \quad (6.73)$$

Each qubit in Q is paired with a qubit in C . A controlled-NOT gate is applied between each pair of qubits such that the resulting state is:

$$\rho_1^{(Q/C)} = \sum_z p_z |z\rangle \langle z|_Q \otimes |z\rangle \langle z|_C. \quad (6.74)$$

The unitary,

$$U_t = \sum_z U_z \otimes |z\rangle \langle z|, \quad (6.75)$$

where U_z is a unitary dependent on the string z , is applied to $\rho_1^{(Q/C)}$. In the context of the previous section, this could be the unitary generated by Eq. 6.70 with $H_b = -\alpha |z\rangle \langle z|$ or $H_b^{(l)}$ with ground state $|z\rangle$. The resulting state after application of U_t is

$$\rho_2^{(Q/C)} = \sum_z p_z U_z |z\rangle \langle z|_Q U_z^\dagger \otimes |z\rangle \langle z|_C. \quad (6.76)$$

This process is now unitary and therefore isolated. The diagonal entropy of the initial state is S_0 . At the end of the process, the diagonal entropy

cannot decrease as a result. The diagonal entropy of the C register (i.e. the diagonal entropy of $\text{Tr}_Q \left(\rho_2^{(Q/C)} \right)$) is exactly S_0 . Hence, the diagonal entropy of the Q register must be greater than or equal to zero. For the adiabatic cycle discussed in the beginning of the section, the second law of thermodynamics is not violated, since all the entropy has been moved from Q to C . Only by neglecting C does it appear that the entropy of the system is decreasing. BQA is making use of a bath of classical bits C prepared in a low entropy configuration to hopefully reduce the entropy of Q . Any statement that relies on the system being isolated requires the inclusion of C .

6.5 Conclusion

Planck's principle is a physically reasonable assumption in a broad range of circumstances. In this work, we have used it to motivate continuous-time quantum optimisation (CTQO) without appealing to the adiabatic theorem. Pure state statistical physics provides a novel way of investigating CTQO without appealing to the minutiae of the energy spectrum that would typically be inaccessible for large quantum systems. Statistical physics also provides new ways of discussing and reasoning about CTQO, by invoking (for example) temperature.

Planck's principle motivates forward quantum annealing (QA) with monotonic schedules, but raises questions over cyclic processes (such as reverse quantum annealing). We demonstrated that $\langle H_p \rangle$ can only improve under monotonic quenching, and that the addition of more stages will lower the energy and is likely to result in a better performance. Applying the eigenstate thermalisation hypothesis to the continuum limit opens up the possibility of analysing a new set of equations. These equations are particularly amenable when an ansatz for the partition function is made. There is scope to extend these equations to capture transitions (or equivalently heating).

With improved knowledge of how systems behave away from the adia-

6. STATISTICAL-MECHANICS APPLIED TO CONTINUOUS-TIME QUANTUM OPTIMISATION

batic limit, it may be possible to design better heuristic continuous-time quantum algorithms for optimisation. Studying CTQO is to understand the limitations of Planck's principle, or equivalently Kelvin's statement of the second law of thermodynamics, within the quantum regime.

Chapter 7

Rapid Hamiltonians for combinatorial optimisation problems

In the absence of mature hardware, CTQO has relied on the adiabatic theorem as a guiding design principle. Throughout this dissertation, we have used pure-state statistical physics to understand these quantum algorithms as they move away from the adiabatic limit. In this chapter, we explore a novel design tenet, inspired by Hamiltonians for optimal state-transfer. These are Hamiltonians that transfer the system from the initial state to the final state in the shortest possible time. In this chapter, we use these Hamiltonians as our underlying design principle. Such Hamiltonians typically require knowledge of the final state (i.e., the solution to the optimisation problem). In the absence of this information, we therefore focus on how the behaviour of these Hamiltonians might be approximated to find approximate solutions to optimisation problems. The result is a rapid continuous-time approach, with a single variational parameter. The next section details Hamiltonians for optimal state transfer and the approximated Hamiltonians considered in the rest of the chapter. Throughout this chapter, we work within the QA-framework, where the aim is to transfer from the ground state of H_d to low energy eigenstates of H_p .

7.1 Hamiltonian design

7.1.1 The optimal Hamiltonian

The aim of optimal state-transfer is to find a Hamiltonian that transfers the system from an initial state (i.e., $|\psi_i\rangle$) to a known final state (i.e., $|\psi_f\rangle$) in the shortest possible time. We shall refer to this Hamiltonian as the optimal Hamiltonian (although it is by no means unique).

One notable approach to finding the optimal Hamiltonian comes from Nielsen *et al.* [162, 163, 164] who investigated the use of differential geometry, at the level of unitaries, to find geodesics connecting the identity to the desired unitary. The length of the geodesic was linked to the computational complexity of the problem [163]. A second approach comes from Carnali *et al.* Inspired by the brachistochrone problem, they developed a variational approach at both the level of state-vectors [165] and unitaries [166]. The quantum brachistochrone problem has generated a considerable amount of literature and interest [167, 168, 169, 170, 171, 172].

Both approaches allow for constraints to be imposed on the Hamiltonian. Both concluded that if the only constraint is on the total energy of the Hamiltonian, the optimal Hamiltonian is constant in time. In the rest of this section, we outline the geometric argument put forward by Brody *et al.* [173] to find the optimal Hamiltonian in this case. The reader is invited to refer to the original work for the explicit details.

The optimal Hamiltonian will generate evolution in a straight line in the (complex-projective) space in which the states live. Intuitively, a line in this space between $|\psi_i\rangle$ and $|\psi_f\rangle$ consists of superpositions of the two states. Hence, the line in the complex-projective space can be represented on the Bloch sphere.

If, without loss of generality, $|\psi_i\rangle$ and $|\psi_f\rangle$ are placed in the traditional $z - x$ plane of the Bloch sphere, it is clear that the optimal Hamiltonian generates rotations in this plane. The optimal Hamiltonian is then

(reminiscent of the cross-product):

$$H_{\text{opt}} = -i (|\psi_i\rangle \langle \psi_f| - |\psi_f\rangle \langle \psi_i|). \quad (7.1)$$

This can be scaled to meet the condition on the energy of the Hamiltonian. If $|\psi_f\rangle$ and $|\psi_i\rangle$ are not chosen to lie on the $z - x$ plane of the Bloch sphere, i.e. $\text{Im}(\langle \psi_f | \psi_i \rangle) \neq 0$ the Eq. 7.1 becomes:

$$H_{\text{opt}} = -i (e^{-i\varphi} |\psi_i\rangle \langle \psi_f| - e^{i\varphi} |\psi_f\rangle \langle \psi_i|), \quad (7.2)$$

where $\varphi = \arg(\langle \psi_f | \psi_i \rangle)$. Note that $\text{Tr } H_{\text{opt}} = 0$. Throughout this chapter it is assumed that $\varphi = 0$, or equivalently that $e^{-i\varphi}$ has been absorbed into $|\psi_i\rangle$.

It then remains to calculate the time required to transfer between the two states. This will depend on how far apart the states are and how fast the evolution is. This is encapsulated in the Anandan-Aharonov relationship [174]:

$$\frac{ds}{dt} = 2\delta E(t). \quad (7.3)$$

The left-hand-side denotes the speed of the state, $|\psi(t)\rangle$, where ds is the infinitesimal distance between $|\psi(t+dt)\rangle$ and $|\psi(t)\rangle$ ¹. Under evolution by the Schrödinger equation, with Hamiltonian H , the instantaneous speed of the evolution is given by the uncertainty in the energy, $\delta E^2(t) = \langle \psi(t) | H^2(t) | \psi(t) \rangle - \langle \psi(t) | H(t) | \psi(t) \rangle^2$.

Since the optimal Hamiltonian is constant in time, δE can be evaluated using the initial state. Therefore, the time of evolution is:

$$T = \frac{\arccos |\langle \psi_f | \psi_i \rangle|}{\sqrt{1 - |\langle \psi_f | \psi_i \rangle|^2}}. \quad (7.4)$$

In standard QA, the time-varying Hamiltonian interpolates between H_d

¹The distance is measured by the Fubini-study metric, $ds^2 = 4(1 - |\langle \psi(t) | \psi(t+dt) \rangle|^2)$, on the complex-projective space.

and an Ising Hamiltonian H_p , to evolve from the ground-state of H_d to the ground-state of H_p . Hence, in the rest of this chapter, we identify $|\psi_i\rangle$ with the ground state of H_d and $|\psi_f\rangle$ with the ground-state of H_p . Consider the overlap of H_d with Eq. 7.1:

$$\begin{aligned}\text{Tr}\{H_d H_{\text{opt}}\} &= -i \text{Tr}\{H_d (|\psi_i\rangle \langle \psi_f| - |\psi_f\rangle \langle \psi_i|)\} \\ &= E_d^{(0)} \text{Tr}\{H_{\text{opt}}\} \\ &= 0,\end{aligned}$$

where $E_d^{(0)}$ is the ground-state energy of H_d . Similarly,

$$\text{Tr}\{H_p H_{\text{opt}}\} = 0.$$

Eq. 7.1 has no overlap with any of the Hamiltonians typically used in QA. More generally, if M is any operator whose eigenstates include $|\psi_i\rangle$ or $|\psi_f\rangle$, then

$$\text{Tr}\{M H_{\text{opt}}\} = 0.$$

This means Eq. 7.1 has no overlap with other Hamiltonians, such as Hamiltonians consisting of only XX terms. Any Pauli operator that commutes with H_d or H_p has zero overlap with H_{opt} . As a final example, consider the optimal Hamiltonian for Max-cut on a two-regular graph with four qubits, starting in the plus state. The Hamiltonian (up to some scaling factor):

$$\begin{aligned}H_4 &= Z_1 Y_2 + Y_1 Z_2 + Z_2 Y_3 + Y_2 Z_3 + Z_3 Y_4 + Y_3 Z_4 + Z_1 Y_4 + Y_1 Z_4 \\ &\quad - Z_1 Y_3 - Y_1 Z_3 - Z_2 Y_4 - Y_2 Z_4 - Y_1 Z_2 Z_3 Z_4 - Z_1 Y_2 Z_3 Z_4 - Z_1 Z_2 Y_3 Z_4 \\ &\quad - Z_1 Z_2 Z_3 Y_4 + X_1 Y_2 Z_3 + Y_1 X_2 Z_4 - Y_1 X_2 Z_3 - X_1 Y_2 Z_4 + Y_1 Z_2 X_3 \\ &\quad + X_1 Z_2 Y_3 + X_1 Y_3 Z_4 + Y_1 X_3 Z_4 + Y_1 Z_2 X_4 + X_1 Z_3 Y_4 - Y_1 Z_3 X_4 \\ &\quad - X_1 Z_2 Y_4 + Z_1 Y_2 X_3 + X_2 Y_3 Z_4 - Z_1 X_2 Y_3 - Y_2 X_3 Z_4 + Z_1 Y_2 X_4 \\ &\quad + Y_2 Z_3 X_4 + X_2 Z_3 Y_4 + Z_1 X_2 Y_4 + Z_2 Y_3 X_4 + Z_1 X_3 Y_4 - Z_1 Y_3 X_4 \\ &\quad - Z_2 X_3 Y_4 + X_1 X_2 Y_3 Z_4 + Y_1 X_2 X_3 Z_4 - X_1 Y_2 X_3 Z_4 + Y_1 Y_2 Y_3 Z_4\end{aligned}$$

$$\begin{aligned}
& + X_1 Y_2 Z_3 X_4 + X_1 X_2 Z_3 Y_4 - Y_1 X_2 Z_3 X_4 + Y_1 Y_2 Z_3 Y_4 + Y_1 Z_2 X_3 X_4 \\
& + X_1 Z_2 Y_3 X_4 - X_1 Z_2 X_3 Y_4 + Y_1 Z_2 Y_3 Y_4 + Z_1 Y_2 X_3 X_4 + Z_1 X_2 X_3 Y_4 \\
& - Z_1 X_2 Y_3 X_4 + Z_1 Y_2 Y_3 Y_4, \quad (7.5)
\end{aligned}$$

solves the problem in the shortest possible time. This Hamiltonian has a huge number of terms, yet none of them are the ones typically used in QA. Note that every term involves a Y Pauli. Although Eq. 7.5 is clearly not implementable, it raises questions over whether QA uses the correct terms in the Hamiltonian to achieve practical speed-up, especially in NISQ devices.

7.1.2 Adapting the optimal Hamiltonian

In summary, $e^{-iH_{\text{opt}}T} |\psi_i\rangle$ generates the state $|\psi_f\rangle$. The goal of the rest of this chapter is to harness some of the physics behind this expression for computation. To this end, we primarily focus on Hamiltonians which are constant in time. In the style of a variational quantum eigensolver (VQE) [75, 175], we allow T to be a variational parameter that needs to be optimised in our new approach. In this chapter, we select the T that minimises the final value of $\langle H_p \rangle$. This is done numerically by brute force grid search for the Hamiltonians proposed in this chapter. The grid search corresponds to dividing the time-interval $[0, 2\pi]$ into 1000 equally spaced intervals. For QAOA $p = 1$, the grid was 100 by 100 in the interval $\beta \in [0, \pi]$ and $\gamma \in [0, 2\pi]$. By using brute force search, we minimise the effect of the classical optimiser on the quantum algorithm. There is scope to explore other metrics besides $\langle H_p \rangle$ [176, 177]. As T is a variational parameter, the Hamiltonian is only important up to some constant factor. Rewriting Eq. 7.1 up to some constant gives:

$$H_{\text{opt}} \propto \frac{1}{2i} [|\psi_i\rangle \langle \psi_i|, |\psi_f\rangle \langle \psi_f|], \quad (7.6)$$

assuming $|\psi_i\rangle$ and $|\psi_f\rangle$ have a non-zero overlap (this is a given in the standard QA framework). This equation (Eq. 7.6) provides the starting

point for all the Hamiltonians considered in this chapter.

The optimal Hamiltonian (i.e., Eq. 7.6) requires knowledge of the final state. In practice, when attempting to solve an optimisation problem, one doesn't have direct access to $|\psi_f\rangle$. Instead, one has easy access to H_d and H_p . Therefore, we make the pragmatic substitutions $|\psi_i\rangle\langle\psi_i| \rightarrow H_d$ and $|\psi_f\rangle\langle\psi_f| \rightarrow H_p$ into Eq. 7.6

$$\begin{aligned}
H_{\text{opt}} &\propto \frac{1}{2i} [|\psi_i\rangle\langle\psi_i|, |\psi_f\rangle\langle\psi_f|] \\
&\quad \downarrow \qquad \downarrow \\
H_1 &= \frac{1}{2i} [H_d, H_p].
\end{aligned} \tag{7.7}$$

This Hamiltonian is the most amenable to NISQ implementation of all the Hamiltonians considered in this chapter, therefore the bulk of the chapter is devoted to demonstrating its performance. The results can be seen in Sec. 7.2.

The substitutions $|\psi_i\rangle\langle\psi_i| \rightarrow H_d$ and $|\psi_f\rangle\langle\psi_f| \rightarrow H_p$ introduce errors, such that the evolution under H_1 no longer closely follows the evolution under H_{opt} . In Sec. 7.3 we try to correct for this error by adding a new term to the Hamiltonian,

$$H_{1,\text{improved}} = H_1 + H_{\text{QZ}}. \tag{7.8}$$

The proposed form of H_{QZ} is motivated by the quantum Zermelo problem [178, 179, 180, 181].

Finally, in Sec. 7.4 we exploit our knowledge of the initial state and propose the substitution $|\psi_f\rangle\langle\psi_f| \rightarrow f(H_p)$, where $f(\cdot)$ is some real function:

$$\begin{aligned}
 H_{\text{opt}} &\propto \frac{1}{2i} [|\psi_i\rangle\langle\psi_i|, |\psi_f\rangle\langle\psi_f|] \\
 &\quad \downarrow \\
 H_{\psi_i} &= \frac{1}{2i} [|\psi_i\rangle\langle\psi_i|, f(H_p)] .
 \end{aligned} \tag{7.9}$$

7.2 Taking the commutator between the initial and final Hamiltonian

In Sec. 7.1. we motivated the Hamiltonian

$$H_1 = \frac{1}{2i} [H_d, H_p] \tag{7.10}$$

by substituting out the projectors in Eq. 7.6 for easily accessible Hamiltonians. In this section, we explore the effectiveness of these substitutions. We begin by demonstrating that Eq. 7.10 generates the optimal rotation for a single qubit (Sec. 7.2.1). In Sec. 7.2.2 we show that H_1 has the potential to outperform random guessing within the QA-framework. The rest of the section analyses the performance of H_1 on Max-cut on regular and binomial graphs.

7.2.1 The optimal approach for a single qubit

Here we outline a simple geometric argument which shows that H_1 generates the optimal rotation for a single qubit (hence the name H_1). The eigenstates of H_d and H_p can be represented as points on the surface of the Bloch sphere, see Fig. 7.1. Since these points lie in a plane, the aim is to write down a Hamiltonian that generates rotation in this plane. To simplify the calculations, we make use of index notation and the Einstein summation notation convention. For this reason, in this section the k^{th} Pauli matrix is denoted by σ_k , with σ_0 being the identity.

The first step is to construct traceless Hamiltonians with a trace-norm of

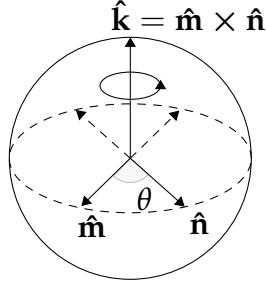


Figure 7.1: The geometric intuition behind finding the Hamiltonian for optimally transferring between the ground-states of H_d and H_p on the Bloch sphere. The vectors $\pm\hat{\mathbf{m}}$ ($\pm\hat{\mathbf{n}}$) are the eigenvectors of H_d (H_p). The aim is to generate a rotation of θ around $\hat{\mathbf{k}}$ to map $\pm\hat{\mathbf{m}}$ to $\pm\hat{\mathbf{n}}$. The handedness of the cross-product takes into account the direction.

one.

$$\tilde{H}_d = \frac{H_d - 1/2 \text{Tr}(H_d) \sigma_0}{\sqrt{\frac{1}{2} \text{Tr}[(H_d - 1/2 \text{Tr}(H_d) \sigma_0)^2]}}, \quad (7.11)$$

and

$$\tilde{H}_p = \frac{H_p - 1/2 \text{Tr}(H_p) \sigma_0}{\frac{1}{2} \sqrt{\text{Tr}[(H_p - 1/2 \text{Tr}(H_p) \sigma_0)^2]}}. \quad (7.12)$$

The Hamiltonians, \tilde{H}_d and \tilde{H}_p , have both the same eigenvectors, and ordering in terms of energy, as H_d and H_p . Expanding \tilde{H}_d and \tilde{H}_p in terms of Pauli matrices gives: $\tilde{H}_d = \hat{\mathbf{m}} \cdot \vec{\sigma}$ and $\tilde{H}_p = \hat{\mathbf{n}} \cdot \vec{\sigma}$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and, $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are real vectors with Euclidean norm of one.

The eigenvectors of \tilde{H}_d and \tilde{H}_p correspond to $\pm\hat{\mathbf{m}}$ and $\pm\hat{\mathbf{n}}$ in the Bloch sphere representation. Ignoring the trivial case, when $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are parallel, the two vectors define a plane. The vector $\hat{\mathbf{k}} = \hat{\mathbf{m}} \times \hat{\mathbf{n}}$ is perpendicular

to the plane and $\hat{k} \cdot \vec{\sigma}$ generates rotations in the \hat{m}, \hat{n} plane. Note that:

$$\begin{aligned} \frac{1}{2i} [\tilde{H}_d, \tilde{H}_p] &= \frac{1}{2i} [m_i \sigma_i, n_j \sigma_j] \\ &= \frac{1}{2i} m_i n_j [\sigma_i, \sigma_j] \\ &= m_i n_j \varepsilon_{ijk} \sigma_k \\ &= \varepsilon_{kij} m_i n_j \sigma_k \\ &= (\hat{m} \times \hat{n}) \cdot \vec{\sigma}, \end{aligned}$$

where ε_{ijk} is the Levi-Civita tensor [182].

It is clear that $[\tilde{H}_d, \tilde{H}_p] / 2i$ generates a rotation in the plane spanned by \hat{m} and \hat{n} . This result also follows trivially from Eq. 7.1 for the single qubit case. The second step is to calculate the angle, θ , between the respective ground states (which is the same as the angle between the excited states). This can be deduced from the overlap between \tilde{H}_d and \tilde{H}_p :

$$\begin{aligned} \frac{1}{2} \text{Tr}(\tilde{H}_d \tilde{H}_p) &= \frac{1}{2} \text{Tr}(m_i n_j \sigma_i \sigma_j) \\ &= \frac{1}{2} m_i n_j \text{Tr}(\delta_{ij} I + i \varepsilon_{ijk} \sigma_k) \\ &= m_i n_i \\ &= \cos(\theta). \end{aligned}$$

Relating this angle to a time can be done by using the Anandan-Aharonov relation:

$$\frac{d\theta}{dt} = 2\delta E(t), \quad (7.13)$$

where $\delta E(t) = \sqrt{\langle \psi(t) | H(t)^2 | \psi(t) \rangle - \langle \psi(t) | H(t) | \psi(t) \rangle^2}$. The Hamiltonian being considered is constant in time. Thus, $\delta E(t)$ can be calculated using the initial state (i.e., $\delta E(0) := \delta E$). Hence, we can deduce the time of evolution as:

$$T = \frac{\arccos [\text{Tr}(\tilde{H}_d \tilde{H}_p) / 2]}{2\delta E}. \quad (7.14)$$

Finally, the Hamiltonians, H_0 and H_1 , can be replaced by the original Hamiltonians. The time-optimal Hamiltonian for a single-qubit is:

$$H_1 = \frac{1}{2i} [H_d, H_p]. \quad (7.15)$$

The corresponding time is:

$$T = \frac{\arccos \left[\frac{1}{2} \text{Tr} (\tilde{H}_d \tilde{H}_p) \right]}{2\delta E}, \quad (7.16)$$

where δE is the uncertainty in energy corresponding to H_1 , and \tilde{H}_d, \tilde{H}_p are as defined earlier in this section.

It remains an open question as to how much of this geometric intuition in three dimensions can be mapped onto higher dimensional problems. Promisingly, the final operations (commutator, trace) are well-defined outside three dimensions. Having established H_1 as the optimal Hamiltonian for a single qubit, the next section investigates its performance on larger problems.

7.2.2 Application to larger problems

Outperforming random guessing for short times

In this section, we demonstrate that H_1 can always do better than random guessing within the QA-framework. Starting with the time-dependent Schrödinger equation:

$$|\dot{\psi}(t)\rangle = -iH_1 |\psi(t)\rangle,$$

we expand $|\psi(t)\rangle$ in terms of the eigenbasis of H_p , so $|\psi(t)\rangle = \sum_k c_k(t) |k\rangle$ where $|k\rangle$ are the eigenvectors of H_p with associated eigenvalue E_k . The eigenvalues are ordered such that $E_0 \leq E_1 \leq E_2 \dots$

Substituting this into the Schrödinger equation gives:

$$\begin{aligned}\sum_k \dot{c}_k(t) |k\rangle &= -\frac{i}{2i} \sum_k c_k(t) (H_d H_p - H_p H_d) |k\rangle \\ &= -\frac{1}{2} \sum_k c_k(t) E_k (H_d - H_p H_d) |k\rangle\end{aligned}$$

Acting with $\langle j|$ on each side, to find $\dot{c}_j(t)$, gives:

$$\dot{c}_j(t) = -\frac{1}{2} \sum_k c_k(t) \underbrace{(E_k - E_j)}_{\text{"Velocity"}} \overbrace{\langle j| H_d |k\rangle}^{\text{How the basis states of } H_p \text{ are connected}}. \quad (7.17)$$

In the standard QA-framework $H_d = -\sum_k^n X_k$, $c_k(0) = 1/\sqrt{2^n}$, for all k , and the basis states (e.g. $|k\rangle$), correspond to computational basis states. Accordingly, H_1 connects computational basis states which are a Hamming-distance of one away.

The difference in energy of the computational basis states intuitively provides something akin to a velocity, with greater rates of change between states which are further apart in energy.

Focusing on the derivative of the ground-state amplitude at $t = 0$ we have:

$$\dot{c}_0(0) = -\frac{1}{2} \sum_k \underbrace{c_k(0)}_{>0} \underbrace{(E_k - E_0)}_{\geq 0} \underbrace{\langle 0| - \sum_j X_j |k\rangle}_{=0 \text{ or } -1}. \quad (7.18)$$

Hence $\dot{c}_0(0) \geq 0$, with equality if all states in a Hamming distance of one have the same energy as $|0\rangle$. In this case, the above logic can be repeated for these states. Hence, at $t = 0$, the ground state amplitude is increasing, meaning that H_1 can do better than random guessing by measuring on short times. This is evidence that H_1 is capturing something of the optimal Hamiltonian for short times. Indeed, for short times, all the amplitudes flow from higher energy states to lower energy states.

The above logic can be extended to the case where H_d is any stoquastic Hamiltonian in the computational basis [48, 183] and $|\psi_i\rangle$ the corresponding ground state. That is to say, we require H_d to have non-positive off-diagonal elements in the computational basis (i.e. be stoquastic) and as a consequence we can write the ground-state of H_d with real non-negative amplitudes [48]. Consequently, for any stoquastic choice of H_d the ground-state amplitude is increasing at $t = 0$ and can do better than the initial value of c_0 at short times.

We could also take a generalised version of H_1 :

$$H_{1,gen} = \frac{1}{2i} [f(H_d), g(H_p)], \quad (7.19)$$

where f and g are real functions. Eq. 7.17 becomes:

$$\dot{c}_j(t) = -\frac{1}{2} \sum_k c_k(t) (g(E_k) - g(E_j)) \langle j | f(H_d) | k \rangle. \quad (7.20)$$

The function acting on H_d (i.e., $f(\cdot)$) controls how the computational basis states are connected, while the function acting on H_p (i.e., $g(\cdot)$) controls the velocity between computational basis states. If $f(\cdot)$ is the identity and H_d stoquastic, then any monotonic function for $g(\cdot)$ (e.g., H_p^3 , H_p^5 , $\exp\{H_p\}, \dots$) will do better than $c_0(0)$ for short t . Taking H_d to be the transverse-field Hamiltonian, that is better than random-guessing.

The above analysis demonstrates that H_1 has potential for tackling generic problems within the QA-framework. The next sections apply H_1 to specific examples in an attempt to quantify the success of this approach. For the rest of this chapter, we take H_d to be the transverse-field Hamiltonian.

Max-cut on two-regular graphs

Here we study the performance of H_1 on Max-cut with two-regular graphs. The explicit form of H_1 is:

$$H_1 = \sum_{j=1}^n (Y_j Z_{j+1} + Z_j Y_{j+1}). \quad (7.21)$$

This can be solved analytically by mapping the problem onto free fermions via the Jordan-Wigner transformation. The end goal is to map $H_1 = \sum_j Y_j Z_{j+1} + Z_j Y_{j+1}$, $H_d = -\sum_j X_j$, and $H_p = \sum_j Z_j Z_{j+1}$ onto fermionic operators. Applying a Fourier Transform then decouples the Hamiltonians into pseudo-spins that can be easily studied. The notation and method closely follow the calculation presented in [80].

The first step is to introduce spin-raising and -lowering operators:

$$S_j^+ = \frac{1}{2} (Y_j + iZ_j), \quad (7.22)$$

$$S_j^- = \frac{1}{2} (Y_j - iZ_j). \quad (7.23)$$

In terms of these new operators:

$$H_d = \sum_j I - 2S_j^+ S_j^-, \quad (7.24)$$

$$H_p = \sum_j S_j^+ S_{j+1}^- + S_j^- S_{j+1}^+ - S_j^+ S_{j+1}^+ - S_j^- S_{j+1}^-, \quad (7.25)$$

$$H_1 = 2i \sum_j S_j^- S_{j+1}^- - S_j^+ S_{j+1}^+. \quad (7.26)$$

The Jordan-Wigner transformation can now be applied to map these spin operators onto fermionic operators, a_j and a_j^\dagger , where:

$$a_j = S_j^- e^{-i\varphi_j}, \quad (7.27)$$

$$a_j^\dagger = S_j^+ e^{i\varphi_j}, \quad (7.28)$$

and $\varphi_j = \pi \sum_{j' < j} a_{j'}^\dagger a_{j'}$. The fermionic operators obey the standard anti-commutation relationship for fermionic operators (i.e., $\{a_j^\dagger, a_k\} = \delta_{j,k}$). In terms of the fermionic operators, the Hamiltonians are:

$$H_d = \sum_{j=1}^n I - 2a_j^\dagger a_j, \quad (7.29)$$

$$\begin{aligned} H_p = \sum_{j=1}^{n-1} a_j^\dagger a_{j+1} - a_j a_{j+1}^\dagger - a_j^\dagger a_{j+1}^\dagger + a_j a_{j+1} \\ + G \left(-a_n^\dagger a_1 + a_n a_1^\dagger + a_n^\dagger a_1^\dagger - a_n a_1 \right), \end{aligned} \quad (7.30)$$

and

$$H_1 = -2i \sum_{j=1}^{n-1} a_j^\dagger a_{j+1}^\dagger + a_j a_{j+1} + 2iG \left(a_n^\dagger a_1^\dagger + a_n a_1 \right), \quad (7.31)$$

where $G = e^{i\pi \sum_{j=1}^n a_j^\dagger a_j}$. For even n , $G = 1$ (anti-periodic boundary conditions — ABC) and for odd n , $G = -1$ (periodic boundary conditions — PBC).

We now apply a Fourier Transform with appropriate p , such that $e^{ipn} = 1$ for PBC and $e^{ipn} = -1$ for ABC, such that

$$c_p = \frac{1}{\sqrt{n}} \sum_j e^{ipj} a_j. \quad (7.32)$$

The Hamiltonians in this new basis are:

$$H_d = \sum_{k=0}^{n-1} I - 2c_k^\dagger c_k \quad (7.33)$$

$$\begin{aligned} H_p = 2 \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \cos \theta_k \left(c_k^\dagger c_k + c_{-k}^\dagger c_{-k} \right) + i \sin \theta_k \left(c_k c_{-k} + c_k^\dagger c_{-k}^\dagger \right) + H_{f,0} \end{aligned} \quad (7.34)$$

and

$$H_1 = 4 \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \sin \theta_k \left(c_k^\dagger c_{-k}^\dagger - c_k c_{-k} \right), \quad (7.35)$$

where for odd n :

$$\begin{aligned} \theta_k &= \frac{2\pi k}{n} \\ H_{f,0} &= -2c_0^\dagger c_0 \\ c_{-k} &= c_{n-k}, \end{aligned}$$

and for even n :

$$\begin{aligned} \theta_k &= \frac{(2k+1)\pi}{n} \\ H_{f,0} &= 0 \\ c_{-k} &= c_{n-1-k}. \end{aligned}$$

These Hamiltonians couple the vacuum state, $|\emptyset\rangle$, with doubly-excited states with opposite momentum, e.g., $c_k^\dagger c_{-k}^\dagger |\emptyset\rangle$. Therefore, we can express the Hamiltonians as pseudo-spins,

$$H_* = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} H_{*,k}, \quad (7.36)$$

with $*$ = $i, f, 1$ and:

$$H_{i,k} = -2Z \quad (7.37)$$

$$H_{f,k} = 2 \cos \theta_k Z - 2 \sin \theta_k Y \quad (7.38)$$

$$H_{1,k} = 4 \sin \theta_k X, \quad (7.39)$$

except for $H_{*,0}$ for $*$ = i, f , for odd n , which is half of the above expressions. The observation that $H_{1,k}$ generates rotations in the plane spanned by the eigenvectors of $H_{f,k}$ and $H_{i,k}$ provides further evidence that H_1 is capturing some of H_{opt} .

The initial state for each pseudo-spin is the ground-state of $-Z$. Calculating the evolution for each pseudo-spin gives:

$$\langle H_p \rangle = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} F_k, \quad (7.40)$$

with $F_0 = 1$ for odd n , otherwise:

$$F_k = 2 \cos \theta_k \cos (8 \sin \theta_k t) - 2 \sin \theta_k \sin (8 \sin \theta_k t). \quad (7.41)$$

The ground state probability is given by:

$$P_{gs} = \prod_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} G_k. \quad (7.42)$$

Again $G_0 = 1$ for odd n , otherwise:

$$G_k = \frac{1}{2} (1 - \cos \theta_k \cos (8 \sin \theta_k t) + \sin \theta_k \sin (8 \sin \theta_k t)). \quad (7.43)$$

This completes the analytical work used to analyse the performance of H_1 for Max-cut on two-regular graphs. It remains to find the optimal time to minimise $\langle H_p \rangle$ – this was done numerically.

A time-domain plot for the approximation ratio and ground-state probability is shown in Fig. 7.2 for 400 qubits. The peak in approximation ratio corresponds to the optimal time. As expected from the previous section (Sec. 7.2.2) the approximation ratio is increasing at $t = 0$. There is a clear peak in ground-state probability at a time of $t \approx 0.275$. This peak in ground-state probability remains present for larger problem sizes too. The peak also occurs at a later time than the peak in approximation ratio. Further insight into this phenomenon may be found in Sec. 7.4.

The key result of this section is shown in Fig. 7.3a, showing the optimal approximation ratio versus problem size for even numbers of qubits

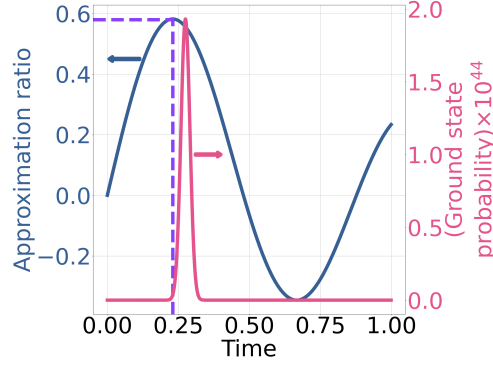
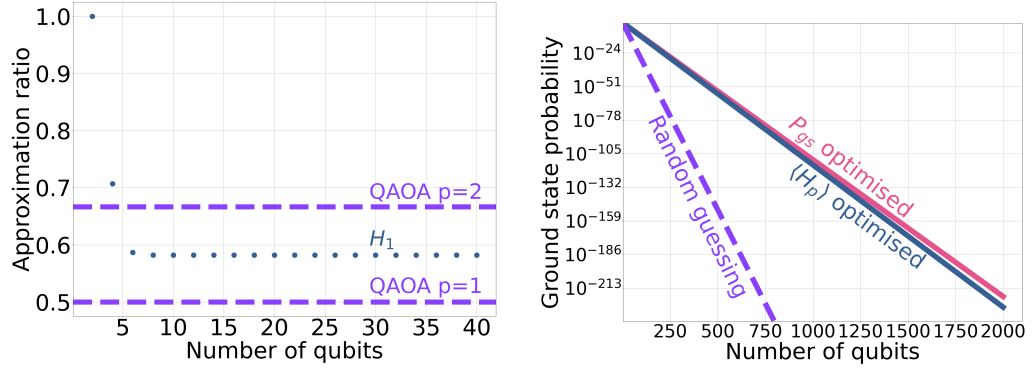


Figure 7.2: A time-domain plot of the ground-state probability (in pink, corresponding to the right-hand side y-axis) and approximation ratio (in blue, corresponding to the left-hand side y-axis) for H_1 applied to Max-cut on a 2-regular graph with 400 qubits. Random guessing corresponds to a ground-state probability of $2^{-399} \approx 10^{-120}$. The dashed purple line shows the location of the optimal time, corresponding to the maximum in approximation ratio.



(a) The approximation ratio for H_1 applied to Max-cut on two-regular graphs (blue dots). The dashed lines show the performance of QAOA for this problem when $p < n/2$ [14, 81]. The corresponding optimal times can be found in Fig. 7.4. The approximation ratio for H_1 freezes out at 0.5819, with a corresponding freeze time of 0.2301.

(b) The ground-state probability for different problem sizes under the evolution of H_1 . The blue line shows the ground-state probability for times that maximise the approximation ratio. The optimised ground-state probability is shown in pink. The dashed purple line shows the probability of randomly guessing the ground-state.

Figure 7.3: Performance of H_1 on Max-cut with 2-regular graphs. Only even numbers of qubits are plotted.

7. RAPID HAMILTONIANS FOR COMBINATORIAL OPTIMISATION PROBLEMS

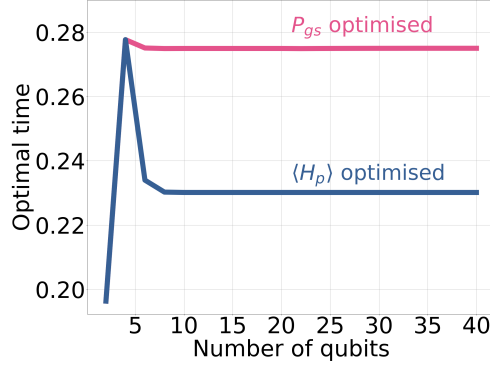
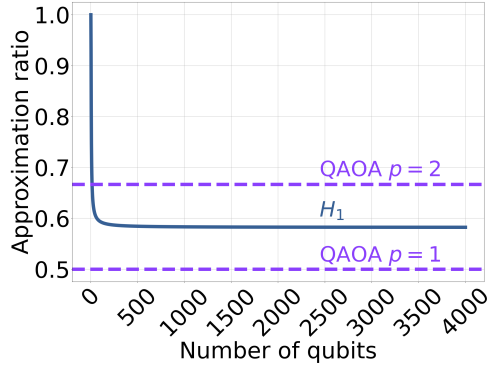
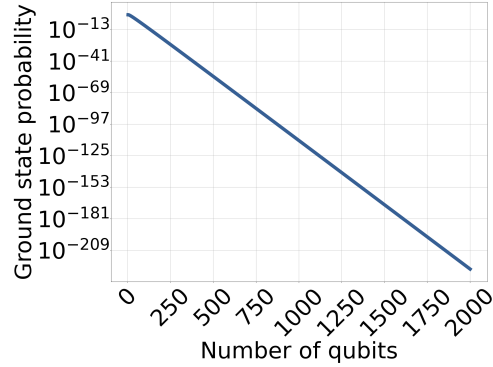


Figure 7.4: The optimal times for Max-cut on two-regular graphs. The blue (pink) line shows the time that optimises the approximation ratio (ground state probability). Even qubit numbers only.



(a) The approximation ratio for H_1 on Max-cut with two-regular graphs, compared to the performance of QAOA. The approximation ratio freezes out at 0.5830 — marginally higher than the even case (see Fig. 7.3a). The time also freezes out at 0.2301 (not shown).



(b) The ground-state probability for different problem sizes under the evolution of H_1 . The system is measured at times that maximises the approximation ratio. Like the even qubit number case (Fig. 7.3b), the ground-state probability scales exponentially. Fitting to the data, the ground-state probability scales as $1.64 \times 2^{-0.33n}$ when $\langle H_p \rangle$ is optimised. The ground-state probability scales as $1.61 \times 2^{-0.32n}$ when the ground-state probability is optimised.

Figure 7.5: The performance of H_1 on Max-cut with two-regular graphs on an odd number of qubits

only. The corresponding plot for odd numbers of qubits can be found in Fig. 7.5a. Notably, the approximation ratio saturates for large problem sizes, achieving an approximation ratio of 0.5819 for even qubit numbers, compared to 0.5 for QAOA $p = 1$ for even qubit numbers. This is despite QAOA $p = 1$ having two variational parameters, compared to the single variational parameter for H_1 . This behaviour is suggestive of H_1 optimising locally, since its approximation ratio is largely independent of problem size.

Despite Max-cut on two-regular graphs being an easy problem, the ground-state probability scales exponentially with problem size (Fig. 7.3b for even qubit numbers and Fig. 7.5b for odd qubit numbers). This is not necessarily a problem, as we present H_1 as an approximate approach only. The ground-state probability scaling is also better than Grover-like scaling [5]. Optimising the performance to give the best ground-state probability provides a small gain in performance but does not change the overall exponential scaling, shown for even qubit numbers only. The optimal times for both approximation ratio and ground-state probability can be found in Fig. 7.4 for even qubit numbers. Both times freeze out at constant values for problem sizes greater than 10 qubits.

We have demonstrated with Max-cut on two-regular graphs, at large problem sizes, that H_1 can provide a better approximation ratio than QAOA $p = 1$. We explore this comparison with QAOA $p = 1$ further in Sec. 7.2.3.

Performance on Max-cut problems with three-regular graphs

By exploiting locality in QA with short run-times, Braida *et al.* [46, 65] were able to prove a lower bound on QA on Max-cut with three-regular graphs. Here we apply this approach to H_1 . We find that that H_1 finds at least 0.6003 times the best cut. Hence, H_1 has a marginally better worst-case than QA (which is 0.5933 times the best cut [46]), when this method is applied. For details of the method the reader is referred to [46], the rest of this section provides the explicit details of this computation.

First, we demonstrate how the bound was derived in [46]. The goal of this approach is to estimate the expectation value of a local observable O_L by simulating part of the system. For this to be a useful estimate it is necessary to quantify the error in doing so, that is to calculate:

$$\varepsilon = \left| \langle \psi_i | U^\dagger(t) O_L U(t) | \psi_i \rangle - \langle \psi_i | U_L^\dagger(t) O_L U_L(t) | \psi_i \rangle \right|, \quad (7.44)$$

where $|\psi_i\rangle$ is the initial state of the system, $U(t)$ is the global unitary evolution of the system and $U_L(t)$ the local unitary we wish to simulate. The error is bounded by,

$$\varepsilon \leq \|U^\dagger(t) O_L U(t) - U_L^\dagger(t) O_L U_L(t)\|, \quad (7.45)$$

where $\|\cdot\|$ is the matrix-norm. Adapting the proof from [46] for the case of evolution under a global time-independent Hamiltonian H and local time-independent Hamiltonian H_L gives:

$$\begin{aligned} & \|U^\dagger(t) O_L U(t) - U_L^\dagger(t) O_L U_L(t)\| \\ &= \left\| \int_0^t ds \frac{d}{ds} \left(U^\dagger(s) U_L(s) U_L^\dagger(t) O_L U_L(t) U_L^\dagger(s) U(s) \right) \right\| \end{aligned}$$

Substitute in the Schrödinger equation for the time derivatives to get:

$$\begin{aligned} & \|U^\dagger(t) O_L U(t) - U_L^\dagger(t) O_L U_L(t)\| \\ &= \left\| \int_0^t ds U^\dagger(s) (H - H_L) U_L(s) \tilde{O}_L(t) U_L^\dagger(s) U(s) \right. \\ & \quad \left. + U^\dagger(s) U_L(s) \tilde{O}_L(t) U_L^\dagger(s) (H_L - H) U(s) \right\|, \end{aligned}$$

where $\tilde{O}_L(t) = U_L^\dagger(t) O_L U_L(t)$. Tidying this up with $\Delta H = H - H_L$ gives:

$$\begin{aligned} & \|U^\dagger(t) O_L U(t) - U_L^\dagger(t) O_L U_L(t)\| \\ &= \left\| \int_0^t ds U^\dagger(s) [\Delta H, U_L(s) \tilde{O}_L U(s)] U(s) \right\| \quad (7.46) \end{aligned}$$

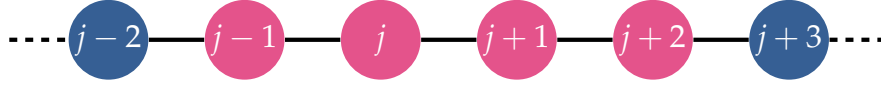


Figure 7.6: A cartoon for calculating the Lieb-Robinson Bound (LRB) on Max-cut with two-regular graphs. Each node represents a qubit, and each edge the interactions between them. To estimate a local expectation value, say $Z_j Z_{j+1}$, the local subgraph in pink is simulated. Bounding the error requires ΔH , corresponding to all the interactions leaving the subgraph. In this case, this corresponds to all the interactions which connect the blue qubits to the pink ones.

Using the triangle inequality,

$$\varepsilon \leq \int_0^t ds \left\| [\Delta H, U_L(s) \tilde{O}_L U(s)] \right\|$$

This is essentially the expression (and proof) given in [46]. For Hamiltonians constant in time, this can further be tidied up to:

$$\varepsilon \leq \int_0^t du \left\| [\Delta H, e^{iH_L u} O_L e^{-iH_L u}] \right\| \quad (7.47)$$

The right-hand-side term in the commutator is a local operator depending only on the local system which we wish to simulate, while ΔH only includes terms in H but not H_L . Therefore, in the context of simulating qubits, the only terms in ΔH that do not commute through are the ones that couple qubits from the local system to qubits outside the system being simulated.

At this point, it is useful to examine a simple system. Take Max-cut on two-regular graphs with H_1 . We wish to estimate $\langle Z_j Z_{j+1} \rangle$, so we estimate the value by simulating the evolution under H_1 for the interactions between the pink qubits shown in Fig. 7.6. Choosing a larger subgraph should make the computation more accurate but will also result in a more difficult computation. Then using Eq. 7.47 we can calculate a bound on the error. Here, ΔH corresponds to the interactions between the pink qubits and the blue qubits. Calculating the local estimate with Lieb-Robinson Bound (LRB) gives Fig. 7.7. The figure also shows the result for the Max-cut for a two-regular graph with 400 qubits.

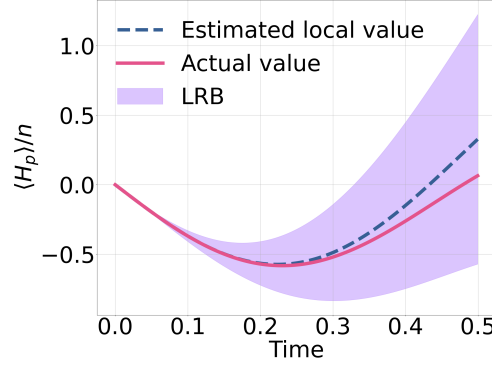


Figure 7.7: LRB applied to Max-cut on two-regular graphs. The dashed blue line shows the local estimate of $\langle Z_j Z_{j+1} \rangle$ and the shaded violet region the allowed region from the LRB. The red line shows the actual value for 400 qubits. The dynamics of H_1 closely resembles the local simulation.

The simulation shows that the bound is not very tight and, except at very short times, is not meaningful. It also demonstrates that H_1 behaves locally, with the full simulation (through the Jordan-Wigner transformation) closely matching the estimate from the local simulation. The LRB on the performance of H_1 for this problem would be the lowest, as a function of time, upper-bound shown in Fig. 7.7.

Having established the idea behind the LRB, we now apply it to find the performance of H_1 on three-regular graphs. To do this, we use

$$H_p = \sum_{(i,j) \in E} Z_i Z_j \quad (7.48)$$

for the graph (or subgraph) $G = (V, E)$ being considered. The minimum value of $\langle H_p \rangle$ is then mapped onto the length of the cut for easier comparison with QAOA and QA.

The performance of H_1 was determined by looking at the three local sub-graphs that can be found in [14, 46]. The error was calculated, taking the worst-case scenario, where each sub-graph has the maximum number of interactions exiting the sub-graph. Using the relative ratios of the sub-graphs [14], the worst-case performance can be calculated. The numerical details can be found in Table 7.1.

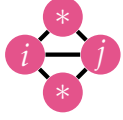
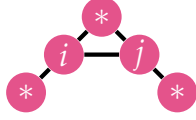
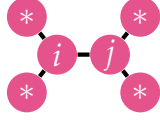
	Subgraph 1	Subgraph 2	Subgraph 3
			
Local estimate of $Z_i Z_j$	-0.2056	-0.2676	-0.3377
Upper estimate from LRB	-0.1333	-0.1652	-0.2007
Cut value	0.5666	0.5826	0.6003

Table 7.1: Numerical details for the LRB applied to H_1 on Max-cut with three-regular graphs. Each column shows, from top to bottom, the local subgraph being simulated; the local estimate of $Z_i Z_j$ to be minimised; the corresponding worst case from the LRB; and the corresponding cut value for this worst case. All of these values are taken at the optimised time of 0.093.

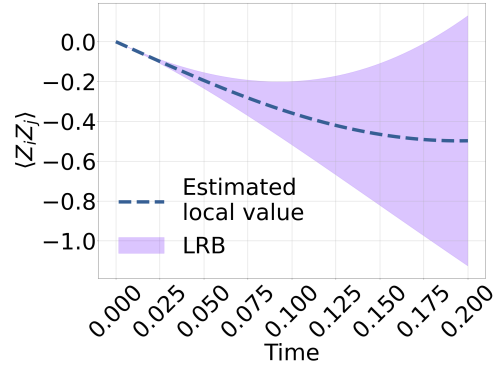
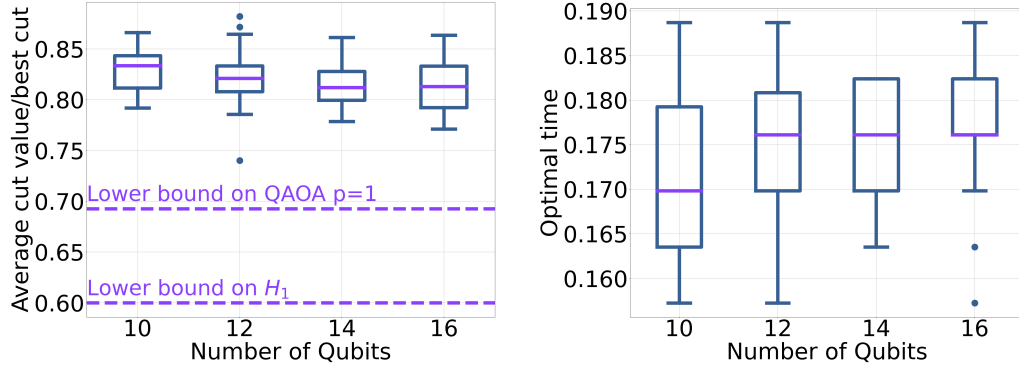


Figure 7.8: The Lieb-Robinson inspired bound for subgraph 3 in Table 7.1. This subgraph dominates the worst-case bound. The LRB has a minimum at around a time of 0.1 while the locally estimated value has a minimum at around a time of 0.2.

Similarly to QAOA and QA [46], the LRB approach suggests that H_1 struggles the most with triangle-free graphs. The LRB, with local estimate, for the triangle-free subgraph can be seen in Fig. 7.8.

The LRB is taken at a time of 0.093, while the minimum for the local estimate occurs at around 0.2. Hence, the LRB is sampling far from what is optimal for the local graph. Since we know the bound is not very tight, it is reasonable to assume that the worst-case performance of H_1 is actually better than the LRB and occurs at a later time (around 0.2). The bound is not necessarily (and unlikely to be) tight. Resorting to numerical simulation gives Fig. 7.9a. Here we can see, for the random

7. RAPID HAMILTONIANS FOR COMBINATORIAL OPTIMISATION PROBLEMS



(a) The y-axis shows the average cut-value from sampling H_1 . The final time has been numerically optimised to give the best approximation ratio.

(b) The optimal time for the three-regular Max-cut instances considered in Fig. 7.9a. The optimal time was found by dividing the interval $[0, 2\pi]$ into 1000 time steps.

Figure 7.9: The performance of H_1 on randomly generated instances of three-regular graphs. For each problem size, 100 instances were generated. After accounting for graph isomorphisms, the number of samples in order of ascending problem size were [15, 46, 87, 97]. Disconnected graphs were allowed.

instances considered, that H_1 never does worse than the QAOA $p = 1$ worst bound. Fig. 7.9a also shows that the approximation ratio of H_1 on three-regular graphs has little dependence on the problem size, again suggesting that it is optimising locally. Directly comparing H_1 and QAOA $p = 1$ gives Fig. 7.10 for three-regular graphs. For all instances H_1 provides a better approximation ratio and in the vast majority of instances in a shorter optimal time. Details on the outliers can be found in Appendix B.2.

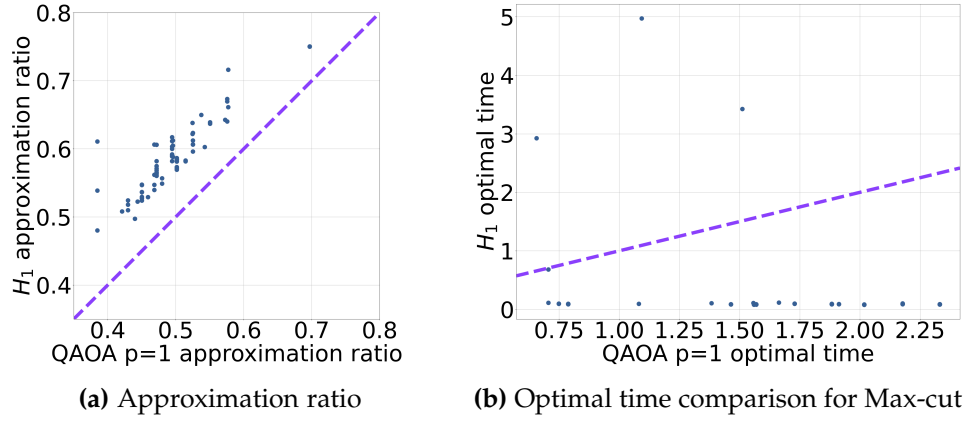


Figure 7.10: Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots) for three-regular graphs. The dashed purple line shows QAOA and H_1 performing the same.

Numerical simulations on random instances of Max-cut

In the previous two sections, we established the performance of H_1 on problems with a high degree of structure, allowing for more analytical investigation. In this section, we explore the performance of H_1 numerically on Max-cut with binomial graphs. Consequently, we are restricted to exploring problem sizes that can be simulated classically.

For each problem size, we consider 100 randomly-generated instances. The results can be seen in Fig. 7.11. For each instance, the time has been numerically optimised to maximise the approximation ratio within the interval $[0, 2\pi)$. From Fig. 7.11 we draw some conclusions from the simulations, with the caveat that either much larger sizes need to be simulated and/or analytic work is required to fully substantiate the claims. On all the problems considered, H_1 performed better than random guessing (which results in an approximation ratio of 0).

We focus first on Fig. 7.11a. There appears to be some evidence that the distribution of approximation ratios is becoming smaller as the problem size is increased. In addition, the approximation ratio tends to a constant value, independent of the problem size. From analysing the regular graphs, it is reasonable to assume that H_1 is optimising locally.

7. RAPID HAMILTONIANS FOR COMBINATORIAL OPTIMISATION PROBLEMS

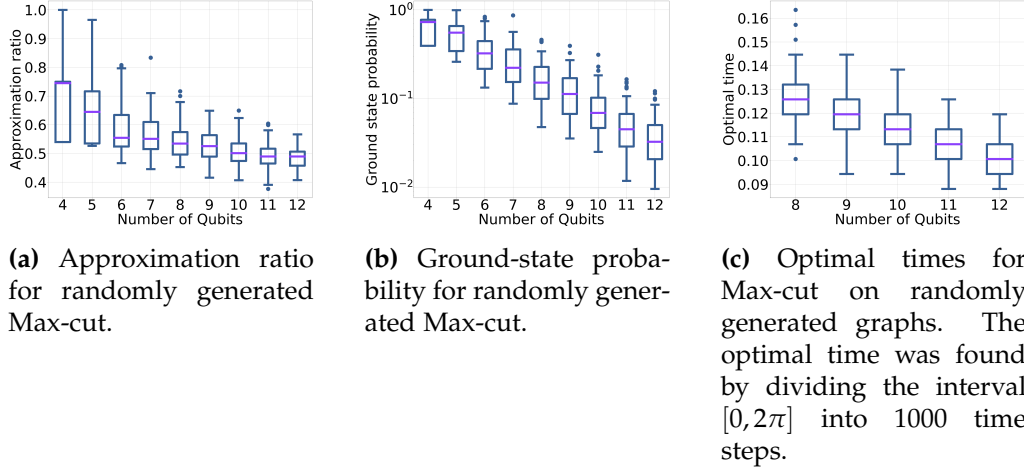


Figure 7.11: Performance of H_1 on 100 randomly-generated instances of Max-cut.

Therefore, we would expect the performance of H_1 to depend on the subgraphs in the problem. If the performance is limited by one subgraph, or a certain combination of subgraphs, then that would explain the constant approximation ratio. As the problem size is increased, the chance of having an atypical combination of subgraphs is likely to decrease, resulting in a smaller distribution.

The ground-state probability shows a clear exponential dependence on problem size (Fig. 7.11b).

Estimating the optimal time

Presenting the application of H_1 as a variational approach begs the question of how to find good initial guesses for the time, T , at which to measure the system. As previously noted, the optimal time corresponds to the maximum approximation-ratio.

For our method, we are not interested in finding the true maximum in approximation ratio. Sampling from a local maximum, close to $t = 0$ is more achievable and reduces the time the system needs to be coherent. From Sec. 7.2.2 we expect the first turning point in approximation ratio after $t = 0$ to be a local maximum. This is shown by Eq. 7.17, with all amplitudes flowing from higher energy states to lower energy states a

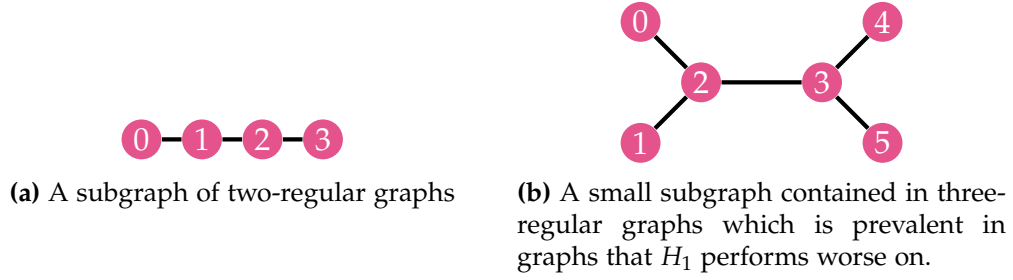


Figure 7.12: Example subgraphs

Hamming distance of one away at $t = 0$. Further to this, throughout the work so far we have seen evidence of H_1 behaving locally. This local behaviour allows us to motivate good initial guesses for H_1 . As H_1 is optimising locally, its performance does not depend on the graph as a whole, but only on subgraphs.

The optimal time for Max-cut on two-regular graphs was $T = 0.23$. Under the assumption that H_1 is behaving locally, we can estimate the optimal times by considering a smaller subgraph. The subgraph in question is shown in Fig. 7.12a. By numerically optimising $\langle Z_1 Z_2 \rangle$ for this subgraph with $H_p = Z_0 Z_1 + Z_1 Z_2 + Z_2 Z_3$, we can find good estimates for the optimal T . Optimising this subgraph, within the interval $T \in [0, 1)$, gives $T = 0.22$. This estimate matches the optimal time well. Choosing a larger subgraph will give a better estimate on the time.

Fig. 7.9b shows the optimal time for the larger instances of three-regular graphs considered in Fig. 7.9a. The range of optimal times varied very little between problem instances and problem sizes, centred around $T = 0.176$. As with the two-regular case, we can examine subgraphs. In this case, we consider the subgraph shown in Fig. 7.12b. This is the subgraph that saturates the Lieb-Robinson bound. Numerically optimising $\langle Z_2 Z_3 \rangle$ for this subgraph with $H_p = Z_0 Z_2 + Z_1 Z_2 + Z_2 Z_3 + Z_3 Z_4 + Z_3 Z_5$ gives a time of $T = 0.19$. This again is a good estimate of the optimal time.

For problems with well understood local structure, such as regular graphs, we have shown that we can exploit this knowledge to provide reasonable estimates of the optimal times. These subgraphs are also of

the same size used in finding the optimal time in QAOA $p = 1$ [14].

For the Max-cut problems in Sec. 7.2.2, the optimal times can be found in Fig. 7.11c. It appears that the optimal time tends to a constant value (or a small range of values), with $T < 1$. The optimal times are clustered together, suggesting good optimal times might be transferable between problem instances. This approach is common within the QAOA literature [184].

So far, we have explored the performance of H_1 . We have demonstrated that H_1 can provide a better approximation ratio for Max-cut on two-regular graphs. The intuition gained by studying QAOA $p = 1$ has been transferable to the understanding of H_1 . In the final part of this section, we make some direct comparisons between QAOA $p = 1$ and H_1 .

7.2.3 Direct numerical comparisons to QAOA $p = 1$

We have established in the previous sections that H_1 operates locally. Calculating the optimal time for sub-graphs the same size as those involved in QAOA $p = 1$ gave good estimates for the optimal time for larger problem sizes. Therefore, it is reasonable to assume that H_1 sees a similar proportion of the graph as QAOA $p = 1$. Both approaches are variational with short run-times too. Since both approaches are using comparable resources, in this section we attempt to compare the two. To make a fair comparison between the two approaches with different problem sizes, we fix the energy of the Hamiltonians in Sec. 7.2.3 to be:

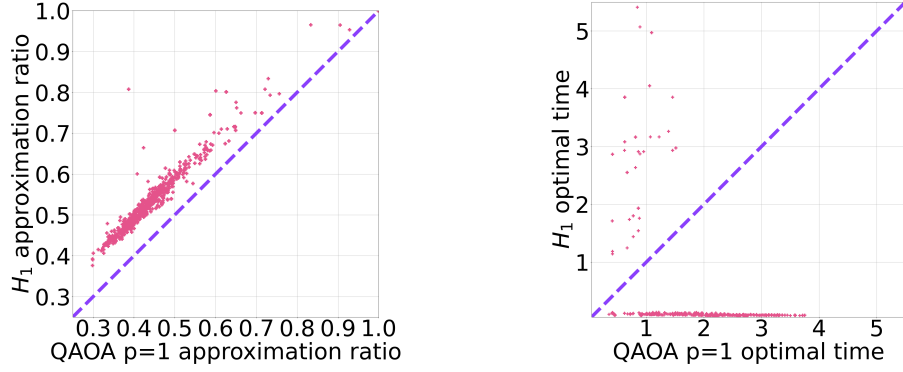
$$\frac{1}{2^n} \text{Tr}\{H_*^2\} = n, \quad (7.49)$$

for $* = d, p, 1$, where n is the number of qubits.

For two-regular graphs, the optimal time for QAOA $p = 1$ is 2.4 times longer than the optimal time for H_1 for large problem sizes², despite

²The optimal time for QAOA is found by optimising the subgraph shown in

providing a poorer approximation ratio.



(a) Approximation ratio comparison for Max-cut on binomial graphs.

(b) Optimal time comparison for Max-cut on binomial graphs.

Figure 7.13: Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots) for the problem instances considered in Sec. 7.2.2. Fig. 7.13a compares approximation ratios. Fig. 7.13b compares the optimal times (i.e. the time that maximises the approximation ratios) of the two approaches. The dashed purple line corresponds to equal performance.

The results comparing H_1 and QAOA $p = 1$, for all the problem instances considered in Sec. 7.2.2, are shown in Fig. 7.13.

The approximation ratios for each approach are largely correlated, suggesting in general that harder problems for QAOA $p = 1$ corresponded to harder problems for H_1 . For all instances considered, H_1 gave a greater than or equal to approximation ratio compared to QAOA $p = 1$ (Fig. 7.13a).

Turning now to the optimal time, H_1 had in the majority of cases the shorter optimal time (Fig. 7.13b). This is shown by the majority of points being below the purple line in Fig. 7.13b, with times less than one. In Appendix B.2 we elaborate further on the exceptions, that is the Max-cut problems that have longer run-times than QAOA $p=1$.

This section has numerically demonstrated that H_1 provides a better approximation ratio than QAOA $p = 1$ in a significantly shorter time

Fig. 7.12a.

for the majority of instances considered, justifying our description of this approach as rapid, which is crucial for NISQ implementation [4]. Given that H_1 tends to provide a better approximation ratio, in a shorter time, with fewer variational parameters, it raises the question — does QAOA $p = 1$, the foundation of any QAOA circuit, make effective use of its afforded resources? Since the run-time for each individual instance, for both shallow-depth QAOA and the H_1 is fixed, it does not determine the scaling of the time to solution. The scaling of the time to solution is determined by the probability of finding the solution (or a desired sufficiently low energy eigenstate).

7.3 An improvement inspired by the Quantum Zermelo problem

7.3.1 The approach

With QAOA it is clear how to get better approximation ratios, that is by increasing p . It is less clear how to do this with H_1 . One suggestion might be to append this Hamiltonian to a QAOA circuit. However, the aim of this chapter is to explore how Hamiltonians for optimal state-transfer can provide a guiding design principle. Therefore, in this section we explore adding another term, motivated by this new design principle, in order to H_1 to improve the approximation ratio.

In Sec. 7.1 we motivated H_1 from the optimal Hamiltonian by making the pragmatic substitutions $|\psi_i\rangle\langle\psi_i| \rightarrow H_d$ and $|\psi_f\rangle\langle\psi_f| \rightarrow H_p$. Subsequently, we demonstrated that H_1 provides a reasonable performance. However, H_1 no longer closely followed the evolution under the optimal Hamiltonian. To partially correct for this error, we add another term to the Hamiltonian:

$$H_{1,improved} = H_1 + H_{QZ}. \quad (7.50)$$

Again, we make use of Hamiltonians for optimal state-transfer to moti-

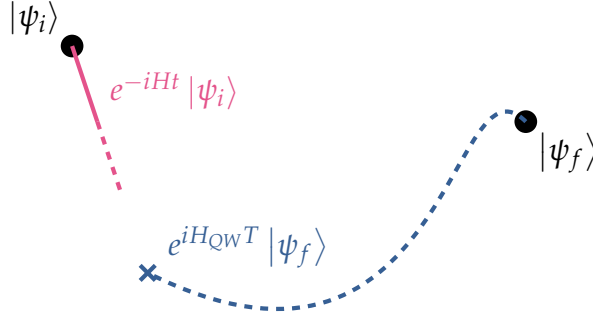


Figure 7.14: A cartoon of the evolution of states in the QZ problem for constant H_{QW} . In the interaction picture, with background Hamiltonian H_{QW} , it appears the final state is moving under the influence of this Hamiltonian. In this frame, Eq. 7.1 can then be applied. It then remains to move out of the interaction picture to get Eq. 7.52.

vate the form of H_{QZ} . Finding the optimal Hamiltonian in the presence of an uncontrollable term in the Hamiltonian is known as the quantum Zermelo (QZ) problem [178, 179, 180, 181].

In the rest of this section, we expand on the details of the QZ problem. From the exact form of the optimal correction, H_{cor} , we then apply a series of approximations so that H_{cor} is time-independent and does not rely on knowledge of $|\psi_f\rangle$. This final Hamiltonian will be H_{QZ} in Eq. 7.50.

The QZ problem, like the quantum brachistochrone problem, asks what is the Hamiltonian that transfers the system from $|\psi_i\rangle$ to the final state $|\psi_f\rangle$ in the shortest possible time. Unlike the quantum brachistochrone problem, part of the Hamiltonian is uncontrollable. In the case of a constant uncontrollable term, the total Hamiltonian can be written as:

$$H_{\text{opt}|QW} = H_{QW} + H_{\text{cor}}(t), \quad (7.51)$$

where H_{QW} is the constant ‘quantum wind’ that cannot be changed and $H_{\text{cor}}(t)$ is the Hamiltonian we are free to vary. Typically, H_{QW} is understood as a noise term [185, 186]. Instead, here we will take H_{QW} to be

H_1 to provide a favourable quantum wind that $H_{\text{cor}}(t)$ can provide an improvement on.

The optimal form of $H_{\text{cor}}(t)$ is (up to some factor) [179]:

$$H_{\text{cor}}(t) = -ie^{-iH_1t} \left(|\psi_i\rangle \langle \psi_f| e^{-iH_1T} - e^{iH_1T} |\psi_f\rangle \langle \psi_i| \right) e^{iH_1t}, \quad (7.52)$$

where t is the time and T is the final time. The motivation for this equation can be found in Fig. 7.14. This Hamiltonian requires knowledge of the final state, so we introduce a series of approximations to make $H_{\text{cor}}(t)$ more amenable for implementation.

Since we know that the optimum evolution under H_1 tends to be short, we make the assumption that the total time T is small. Therefore, we approximate the optimal correction $H_{\text{cor}}(t)$ with

$$H_{\text{cor}}(0) = -i \left(|\psi_i\rangle \langle \psi_f| e^{-iH_1T} - e^{iH_1T} |\psi_f\rangle \langle \psi_i| \right). \quad (7.53)$$

Introducing the commutator structure (Sec. 7.1) with the same pragmatic substitutions as before for H_1 gives:

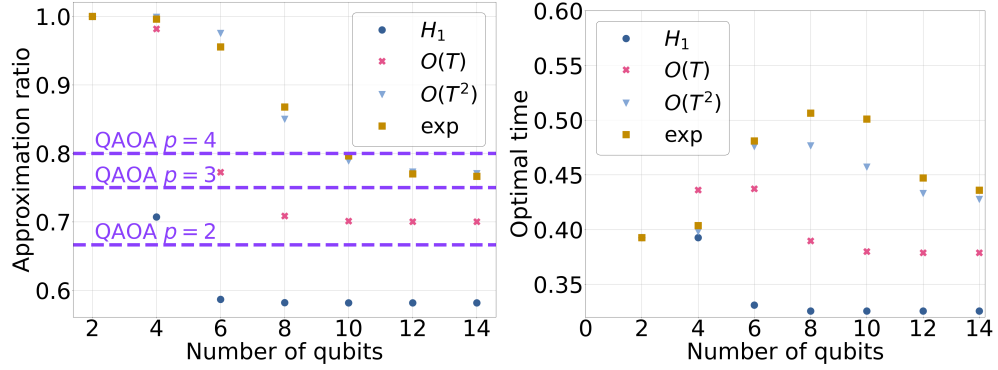
$$H_{QZ} = -i \left[H_d, e^{iH_1T} H_p e^{-iH_1T} \right], \quad (7.54)$$

where we have introduced the subscript QZ to distinguish this Hamiltonian from $H_{\text{cor}}(0)$ prior to the substitutions. Expanding this expression in T gives:

$$H_{QZ} = -i \left[H_d, H_p + iT [H_1, H_p] - T^2 [H_1, [H_1, H_p]] / 2 + \mathcal{O}(T^3) \right]. \quad (7.55)$$

From the QZ problem, we have motivated the form of the correction H_{QZ} in Eq. 7.50. In spite of this, we have no guarantee on its performance — to this end we carry out numerical simulations.

A clear downside to H_{QZ} is the increased complexity, compared to say



(a) Numerically optimised performance (b) The corresponding optimal times for of Eq. 7.55. Each point has been opti- Fig. 7.15a. The norm of each Hamil-
mised in the time interval $[0,0.3]$ by con- tonian, for each problem size, has been
sidering 3000 divisions. fixed so Eq. 7.49 is true, to make compar-
isons fair.

Figure 7.15: The performance of Eq. 7.55 on two-regular graphs. The legend shows the order of T , with 'exp' referring to Eq. 7.54. The dashed lines show the asymptotic performance of QAOA.

QAOA. However, if H_{QZ} is decomposed into a QAOA-style circuit, the single free parameter in H_{QZ} might translate to fewer free parameters in the QAOA circuit, allowing for easier optimisation.

7.3.2 Numerical simulations

Max-cut on two-regular graphs

Here we focus on applying Eq. 7.55 up to various orders in T to Max-cut on two-regular graphs with an even number of qubits. We focus on this problem as it is trivial to scale and the performance of QAOA and H_1 on this problem is well understood.

The results for Max-cut with two-regular graphs can be seen in Fig. 7.15a. Increasing the expansion in T appears to improve the approximation ratio. But the improvement is capped, shown by the data labelled 'exp'. Notably, this approach with a single variational parameter at order T^2 is performing better than QAOA $p = 3$ (with 6 variational parameters) for 10 qubits.

The optimal T for the QZ-inspired Hamiltonians can be seen in Fig. 7.15b. Again, the optimal time for each order in T appears to be tending to a constant value, suggesting this approach is still acting in a local fashion. This is consistent with the approximation ratio plateauing. As we can see, the QZ-inspired approach is still operating in a rapid fashion.

Max-cut on binomial graphs

To complete this section, we examine the performance of the QZ-inspired approach (Eq. 7.55) to the randomly generated instances of Max-cut.

The results for different orders in T for the approximation ratio can be seen in Fig. 7.16a. All the QZ-inspired approaches provide an improvement on the original H_1 Hamiltonian, indexed by 0 in the figures. However, the performance is not monotonically increasing with the order of the expansion. This is not unusual for a Taylor series of an oscillatory function. Consequently, achieving better approximation ratios is not as simple as increasing the order of T . At the same time, this means that it is not necessary to go to high orders in T , with very non-local terms, to achieve a significant gain in performance. For example, in going to first order achieves a substantial improvement.

The optimal times for the QZ-inspired approach can be found in Fig. 7.16b. For clarity, we only show the optimal times for the larger problem instances. As with H_1 the optimal times are clustered for a given order. The lack of dependence on problem size for optimal times and approximation ratios suggests that the QZ-inspired approach is still optimising locally. Compared to the H_1 case, the operators have a larger support. Despite optimising locally, they are optimising less locally than H_1 , hence the increased performance.

Here we have numerically demonstrated that the QZ-inspired approach can provide an improvement over H_1 , suggesting how this new design

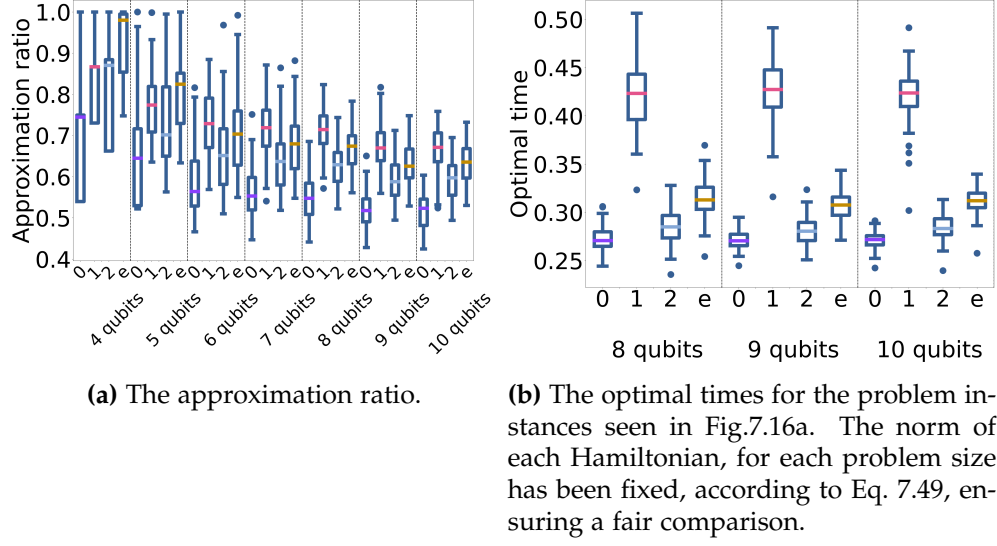


Figure 7.16: The performance of the QZ-inspired approach on 100 random Max-cut instances. The x-axis label refers to the order of T in the expansion of Eq. 7.55, with 0 being H_1 and e referring to the full exponential (i.e. Eq. 7.54)

philosophy might be extended. The numerics also suggest that going to first order may provide the best possible advantage.

7.4 Using knowledge of the initial state

As introduced in Sec. 7.1, in this section, we exploit our knowledge of the initial state and evaluate the performance of

$$H_{\psi_i} = -i [|\psi_i\rangle \langle \psi_i|, f(H_p)] \quad (7.56)$$

within the QA-framework. We take $f(\cdot)$ to be a real function such that:

$$f(H_p) = \sum_k f(E_k) |E_k\rangle \langle E_k|, \quad (7.57)$$

where $|E_k\rangle$ and E_k are the eigenvectors and associated eigenvalues of H_p .

Evolution under H_{ψ_i} can be calculated analytically. By evolving, $|\psi_i\rangle$ the

state:

$$|\omega\rangle = \frac{1}{\sqrt{\text{Tr } f^2(H_p)}} \sum_k f(E_k) |E_k\rangle, \quad (7.58)$$

can be reached. Indeed, H_{ψ_i} will generate linear superpositions of $|\omega\rangle$ and $|\psi_i\rangle$ only. To see this, first rewrite H_{ψ_i} in terms of the eigenstates $|E_k\rangle$ of H_p , with associated eigenenergies E_k :

$$\begin{aligned} H_{\psi_i} &= -i [|\psi_i\rangle \langle \psi_i|, f(H_p)] \\ &= -i \left[|\psi_i\rangle \langle \psi_i|, \sum_k f(E_k) |E_k\rangle \langle E_k| \right] \\ &= -i \left(|\psi_i\rangle \sum_k f(E_k) \langle \psi_i| E_k \rangle \langle E_k| - \sum_k f(E_k) \langle E_k| \psi_i \rangle |E_k\rangle \langle \psi_i| \right). \end{aligned}$$

Let

$$|\omega\rangle = \frac{\sum_k f(E_k) \langle E_k| \psi_i \rangle |E_k\rangle}{\sqrt{\sum_k f^2(E_k) |\langle E_k| \psi_i \rangle|^2}}. \quad (7.59)$$

Apply the Gram-Schmidt procedure to generate an orthonormal basis $\{|\psi_i\rangle, |\omega^\perp\rangle\}$, spanning the same space as $\{|\psi_i\rangle, |\omega\rangle\}$.

$$|\omega^\perp\rangle \propto |\omega\rangle - \langle \psi_i | \omega \rangle |\psi_i\rangle$$

Rearranging for $|\omega\rangle$ gives:

$$|\omega\rangle = \sqrt{1 - |\langle \psi_i | \omega \rangle|^2} |\omega^\perp\rangle + \langle \psi_i | \omega \rangle |\psi_i\rangle, \quad (7.60)$$

substituting this into H_{ψ_i} gives:

$$H_{\psi_i} = -i\eta \left\{ |\psi_i\rangle \left(\langle \omega^\perp | + \langle \omega | \psi_i \rangle \langle \psi_i | \right) - \left(|\omega^\perp\rangle + \langle \psi_i | \omega \rangle |\psi_i\rangle \right) \langle \psi_i | \right\}, \quad (7.61)$$

where

$$\eta = \sqrt{\left(1 - |\langle \psi_i | \omega \rangle|^2\right) \left(\sum_k f^2(E_k) |\langle E_k | \psi_i \rangle|^2\right)}. \quad (7.62)$$

Using the freedom in the global-phase of the wave-function, choose $\langle \omega | \psi_i \rangle$ to be real, giving:

$$H_{\psi_i} = -i\eta \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right). \quad (7.63)$$

From Eq. 7.63 it is clear that H_{ψ_i} evolves $|\psi_i\rangle$ to linear superpositions of $|\psi_i\rangle$ and $|\omega^\perp\rangle$. Again, by using the Anandan-Aharonov relation, we can determine the time to generate $|\omega\rangle$ and $|\omega^\perp\rangle$. Calculating δE :

$$\begin{aligned} \langle H_{\psi_i} \rangle &= -i\eta \langle \psi_i | \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right) | \psi_i \rangle \\ &= 0 \\ \langle H_{\psi_i}^2 \rangle &= \eta^2 \langle \psi_i | G^2 | \psi_i \rangle \\ &= \eta^2 \\ \delta E &= \sqrt{\langle H_{\psi_i}^2 \rangle - \langle H_{\psi_i} \rangle^2} \\ &= \eta. \end{aligned}$$

The distance between $|\psi_i\rangle$ and $|\omega^\perp\rangle$ is

$$\begin{aligned} \theta_{\omega^\perp} &= 2 \arccos \langle \psi_i | \omega^\perp \rangle \\ &= \pi. \end{aligned}$$

The distance between $|\psi_i\rangle$ and $|\omega\rangle$ is

$$\theta_\omega = 2 \arccos |\langle \omega | \psi_i \rangle|.$$

The time then to evolve $|\psi_i\rangle$ to $|\omega^\perp\rangle$ is $t_{\omega^\perp} = \pi/2\eta$ and the time to evolve to $|\omega\rangle$ is $t_\omega = \arccos |\langle \omega | \psi_i \rangle|/\eta$. We can explicitly verify this by exponentiating H_{ψ_i} . Focusing on $G = -i \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right)$, then

$$\begin{aligned} G^2 &= - \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right) \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right) \\ &= \left(|\psi_i\rangle \langle \psi_i| + |\omega^\perp\rangle \langle \omega^\perp| \right). \end{aligned}$$

This is a projector. Calculating G^3 gives:

$$\begin{aligned} G^3 &= -i \left(|\psi_i\rangle \langle \psi_i| + |\omega^\perp\rangle \langle \omega^\perp| \right) \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right) \\ &= -i \left(|\psi_i\rangle \langle \omega^\perp| - |\omega^\perp\rangle \langle \psi_i| \right) \\ &= G. \end{aligned}$$

Let H be a Hamiltonian, where $H^2 = P$, P being a projector. Then substituting H into the power-series for the exponential gives:

$$\begin{aligned} e^{-iHt} &= \sum_{k=0}^{\infty} \frac{(-iHt)^k}{k!} \\ &= \sum_{\text{even } k} \frac{(-iHt)^k}{k!} + \sum_{\text{odd } k} \frac{(-iHt)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(-iHt)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(-iHt)^{2k+1}}{(2k+1)!} \\ &= I + \sum_{k=1}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} P - i \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} PH \\ &= I + (\cos(t) - 1) P - i \sin(t) PH \end{aligned}$$

Therefore,

$$e^{-iH_p t} = I + (\cos(\eta t) - 1) G^2 - i \sin(\eta t) G \quad (7.64)$$

and $|\psi(t)\rangle = e^{-iH_p t} |\psi_i\rangle$,

$$|\psi(t)\rangle = \cos(\eta t) |\psi_i\rangle + \sin(\eta t) |\omega^\perp\rangle. \quad (7.65)$$

Checking the times from the Anandan-Aharonov relationship gives:

$$\begin{aligned} |\psi(\pi/2\eta)\rangle &= \cos(\pi/2) |\psi_i\rangle + \sin(\pi/2) |\omega^\perp\rangle \\ &= |\omega^\perp\rangle \end{aligned}$$

and

$$\begin{aligned}
 & |\psi(\arccos |\langle \psi_i | \omega \rangle| / \eta)\rangle \\
 &= |\langle \psi_i | \omega \rangle| |\psi_i\rangle + \sqrt{1 - |\langle \psi_i | \omega \rangle|^2} |\omega^\perp\rangle \\
 &= |\omega\rangle.
 \end{aligned}$$

Up to now, we have kept the conversation in this section fairly general. Now we apply H_{Ip} to the QA-framework with $|\psi_i\rangle = |+\rangle$ and the $|E_k\rangle$ s corresponding to computational basis states, so $\langle + | E_k \rangle = 1/\sqrt{2^n}$. Simplifying $|\omega\rangle$ gives:

$$\begin{aligned}
 |\omega\rangle &= \frac{1}{\sqrt{\sum_k f^2(E_k) |\langle E_k | + \rangle|^2}} \sum_k f(E_k) \langle E_k | + \rangle |E_k\rangle \\
 &= \frac{1}{\sqrt{\sum_k f^2(E_k)}} \sum_k f(E_k) |E_k\rangle \\
 &= \frac{1}{\sqrt{\text{Tr}\{f^2(H_p)\}}} \sum_k f(E_k) |E_k\rangle
 \end{aligned}$$

and the overlap with the initial state is:

$$\begin{aligned}
 \langle + | \omega \rangle &= \frac{\sqrt{2^n}}{\sqrt{\sum_k f^2(E_k)}} \sum_k \frac{1}{\sqrt{2^n}} f(E_k) \langle + | E_k \rangle \\
 &= \frac{1}{\sqrt{\sum_k f^2(E_k)}} \sum_k \frac{1}{\sqrt{2^n}} f(E_k) \\
 &= \frac{1}{\sqrt{2^n}} \frac{\text{Tr}\{f(H_p)\}}{\sqrt{\text{Tr}\{f^2(H_p)\}}}.
 \end{aligned}$$

To summarise, the Hamiltonian $H_{Ip} = -i[|+\rangle \langle +|, f(H_p)]$ evolves $|+\rangle$ to

$$|\omega\rangle = \frac{1}{\sqrt{\text{Tr} f^2(H_p)}} \sum_k f(E_k) |E_k\rangle, \quad (7.66)$$

in a time

$$T = \frac{\sqrt{2^n} \arccos |\langle +|\omega \rangle|}{\sqrt{\text{Tr}\{f^2(H_p)\} (1 - |\langle +|\omega \rangle|^2)}}, \quad (7.67)$$

where

$$\langle +|\omega \rangle = \frac{1}{\sqrt{2^n}} \frac{\text{Tr}\{f(H_p)\}}{\sqrt{\text{Tr}\{f^2(H_p)\}}}. \quad (7.68)$$

Assuming that the state $|\omega\rangle$ is prepared, then the probability of finding the ground-state is

$$P_{gs} = \frac{g f^2(E_{gs})}{\text{Tr} f^2(H_p)}, \quad (7.69)$$

where g is the ground-state degeneracy and E_{gs} the associated energy. If $f(\cdot)$ is the identity, then $\text{Tr} H_p$ scales approximately as 2^n and E_{gs} might scale with n . Hence, the ground state probability will scale as $\sim n^2/2^n$. Indeed, if $f(H_p) = H_p^m$, where m is some positive integer, then the ground state probability might scale as $\sim n^{2m}/2^n$. This does not include the computational cost of calculating H_p^m , which will be on the order of n^m . This is still an improvement over random guessing, but still with exponential scaling. This exponential scaling is worse than Grover's algorithm [5], or other approaches that provide a better exponential scaling, for example continuous-time quantum walks [22]. This may be of some practical benefit, depending on the computational cost of calculating $f(H_p)$. If $f(\cdot)$ is the projector onto the ground-state, then $P_{gs} = 1$ (as expected).

Calculating the expectation of H_p for $|\omega\rangle$ gives:

$$\langle H_p \rangle = \frac{1}{\text{Tr} f^2(H_p)} \sum_k E_k f^2(E_k). \quad (7.70)$$

Here we can see that $\langle H_p \rangle$ will be dominated by states for which $f^2(E_k)$ is large. If $f(\cdot)$ is the identity this most likely means low energy states and high energy states. Hence, we do not expect a good approximation ratio. This provides some insight into Sec. 7.2.2 where the observed

peak in ground-state probability did not coincide with the optimal approximation ratio.

This approach has the potential to provide a modest practical speed-up with a polynomial prefactor on the hardest problems. However, the success of this approach depends on the (unlikely) feasibility of implementing H_{ψ_i} and $f(H_p)$. It does however provide further evidence of the power of commutators for designing algorithms to tackle optimisation problems.

7.5 Conclusion

Designing quantum algorithms to tackle combinatorial optimisation problems, especially within the NISQ framework, remains a challenge. Many algorithms have used AQO in their inspiration, such as in the choice of Hamiltonians. In this chapter, we have explored using optimal Hamiltonians as a guiding design principle.

With H_1 , the commutator between H_d and H_p , we demonstrated that we can outperform QAOA $p=1$, with fewer resources. The short run-times which do not appear to scale with problem size suggest that this approach is acting locally. An effective Lieb-Robinson bound prevents the information about the problem propagating instantaneously [187, 188]. This helps provide some insights into the performance of H_1 :

- In the local regime, the effective local Hilbert space is smaller than the global Hilbert space, consequently H_1 will be a better approximation of the optimal Hamiltonian. This accounts for why we might expect H_1 to work better on short run-times.
- A local algorithm is unlikely to be able to solve an optimisation problem, as it cannot see the whole graph. It follows that such an approach would have poor scaling of the ground-state probability.

Due to the local nature of H_1 we were able to utilise some analytical tricks to assist the numerical assessment of its performance, allowing for

some guarantee of the performance of the approach on large problem sizes. The techniques used had already been developed or deployed by the continuous-time quantum computing community in the context of QA/QAOA, indicative of the wide applicability of the tools being developed to assess these algorithms.

Local approaches have clear advantages in NISQ-era computations. The short run-times put fewer demands on the coherence times of the device. The local nature can also help mitigate some errors. If, for example, there is a control misspecification in one part of the Hamiltonian this is unlikely to propagate through the whole system and affect the entire computation.

Buoyed by the relative success of utilizing Hamiltonians for optimal state-transfer, we turned to the quantum Zermelo problem to help find improvements. These Hamiltonians comprise a single variational parameter and short run-times, for increased complexity in the Hamiltonian. Again, the saturation of the optimal time suggests that these approaches are still operating locally.

The success of this approach, within the NISQ era, will depend on the feasibility of implementing these Hamiltonians. This might be achieved through decomposition into a product formula [189] for gate based approaches, resulting in a QAOA like circuit. Alternatively, one could attempt to explicitly engineer the interactions involved. Indeed, for exponentially scaling problems, implementing these Hamiltonians for short times could be less challenging than maintaining coherence for exponentially increasing times.

Although the results of this chapter are not fully conclusive, it has shown that by considering Hamiltonians for optimal state-transfer we can develop promising new algorithms. We hope the results presented in this chapter will encourage further work into the success of these Hamiltonians. There is scope for taking this work further. This could include changing the choice of H_d , exploiting our observation that any

stoquastic Hamiltonian can lead to an increase in ground state probability. For H_p we have only explored problems with trivial Ising encodings. There is scope to explore new encodings such as LHZ [190] or Domain-wall [44] encodings. Such encodings will result in different H_1 and presumably distinct dynamics.

Chapter 8

Conclusion

8.1 Summary

This dissertation has explored a range of established continuous-time quantum optimisation approaches through the lens of pure-state statistical physics. It has also explored Hamiltonians for optimal state-transfer as an inspiration for tackling combinatorial optimisation problems.

Chapter 4, motivated by the eigenstate thermalisation hypothesis, used Ansätze for the density of states to optimise continuous-time quantum walks. The free parameters in the Ansätze were extracted from the combinatorial optimisation problem. This allowed for the optimisation of a many-body quantum system where exact numerical diagonalisation could not be achieved. The optimisation of any free parameters in a quantum algorithm needs to be realised in a time efficient manner. In this work we have presented a possible way to do this. Future work might look to extend this to other continuous-time algorithms, as well as better Ansätze.

Having investigated the long-time behaviour of continuous-time quantum walks, Chapter 5 investigated the timescale associated with the dynamics. This was first done by analytically calculating the torsion of the wave function. Combined with an Ansatz for the time evolution, it was found that the timescale associated with continuous-time quantum walks does not scale strongly with the problem size. In practice, the quench might not be instantaneous, with the speed of a ramp limited by

the hardware controls. Understanding the effect of non-instantaneous quenching is important for practical realisations.

Building on the inspiration gained from the eigenstate thermalisation hypothesis, pure-state statistical physics was used to analyse a range of continuous-time quantum optimisation approaches in Chapter 6. Questions were raised over cyclic processes, and it was shown that passive states obey $\langle H_p(0) \rangle \geq \langle H_p(t) \rangle$ for monotonically increasing schedules. This provided a heuristic understanding of multi-stage quantum walks. This work can be extended by making the theory more rigorous by better understanding the limitations and exceptions.

To try to push beyond the conventional adiabatic inspired approaches, Chapter 7 explored using Hamiltonians for optimal state transfer to motivate quantum optimisation approaches. The results were mixed, but it was shown that it could outperform QAOA $p = 1$ with similar resources. Incorporating other techniques from optimal control might help improve on the current proposal. Hopefully, this work will help encourage investigating alternative approaches away from adiabatic inspired approaches.

8.2 Outlook on continuous-time quantum optimisation

Continuous-time quantum optimisation (CTQO) has historically been framed in terms of adiabatic quantum optimisation (AQO) [191, 192]. The aim in AQO is to find the ground state of H_p ; this can be achieved by evolving the system adiabatically from a known ground state. The timescale to remain adiabatic is inversely proportional to a power of the minimum spectral-gap (i.e. the minimum difference in energy between the ground and first excited state over the anneal) [10, 191]. For many problems, this gap is found to close exponentially (also referred to as a vanishing spectral gap), implying an exponentially long runtime [10,

191]. The study of AQO therefore typically involves the study of the spectral gap, converting a dynamical problem to a static problem. An exponentially closing gap is then used to imply a failure of AQO, while a polynomially closing gap implies a success for AQO [48].

Unless one is willing to accept that all hard optimisation problems can be solved efficiently on a quantum computer, then exponentially closing gaps are not necessarily a failure but an expression that there exist hard problems. This is not to say, that a vanishing gap implies hardness, rather that its absence implies that the problem is easy to solve. Therefore, an exponentially vanishing spectral gap is to be expected in some cases,¹ and should not necessarily be used to dismiss AQO.

Despite this, AQO in its current form is unlikely to play a role in practical quantum optimisation. Consider first the case where the gap is exponentially closing. Exponentially increasing runtimes are not feasible, and hence the algorithm will not terminate in a realisable timescale. A decision must be made to terminate the process, resulting in non-adiabatic effects. Alternatively, if the gap is polynomially closing, then AQO can solve the problem efficiently. However, determining the gap for non-trivial problems is typically challenging, with the brute force approach requiring exact diagonalisation throughout the anneal. This means that, in practice, any potential speed-up might be lost, because:

- Verifying that any solution obtained is the ground-state is a computational challenge.
- In the absence of knowledge about the spectral gap, the user might assume that the required runtime is much longer than is required by the adiabatic theorem.

In short, AQO has no role to play in practical CTQO, without connecting the anneal time to accessible quantities that can be obtained prior to the

¹Through a suitably engineered driver Hamiltonian, it is possible to remove the vanishing gap altogether [56]. However, in this case the cost of finding and implementing this driver Hamiltonian should be included in the runtime of the quantum algorithm.

anneal. Even then, this only makes AQO amenable to problems without an exponentially vanishing spectral gap.

AQO focuses on determining the time required to remain adiabatic, despite practical realisations placing restrictions on what times can be achieved. A more sensible question, especially given the finite coherence times achieved in analogue quantum systems, is determining what might be achieved with a finite anneal time T . This time is likely to be much shorter than the anneal-time required to be adiabatic. This is likely to require three major shifts:

1. Accepting and exploiting dynamics.
2. Incorporating more information about the combinatorial optimisation problem.
3. Shifting the focus away from ground-state probabilities.

The rest of this outlook will consider each of these in turn, making links to the content of the dissertation where appropriate.

8.2.1 Accepting and exploiting dynamics

Quantum Annealing (QA) is sometimes conflated with AQO [192], therefore even QA is discussed in terms of vanishing spectral gaps. An exponentially vanishing spectral gap means that the adiabatic path between ground states cannot be practically realised. This does not necessitate a failure of QA or any other CTQO approach. By placing focus on the static problem of the vanishing gap, the role of dynamics is neglected. As discussed, any practical realisation of CTQO will not be adiabatic for at least some problems and therefore dynamics will take place. Theoretically and numerically, studying the dynamics of these strongly interacting spin systems is difficult, but there has been progress [46, 65, 95, 96]. It is not clear yet what insight extensive studies on spectral gaps and AQO will be in this dynamical regime. Ultimately, the spectral gap is being used as a proxy for finding the ground-state of

H_p , despite it being no easier to calculate and only being equivalent in the adiabatic limit. In summary CTQO, including QA, should be accepted as a dynamical algorithm. Given that dynamics is unavoidable, it is important that the role of dynamics is better understood and exploited. Continuous-time quantum walks for optimisation present a step forwards in investigating the role of dynamics in CTQO but more is required to understand the limitations of these approaches.

Given that, as discussed in Chapter 6, multi-stage quantum walks (MSQWs) are approximating thermal states on average, the use of MSQWs as a dynamical approach will depend on a number of factors:

- Can an MSQW approximate a thermal state faster than other classical or quantum methods?
- MSQWs are not perfect Gibbs states. Does the presence of any athermal behaviour provide any advantage?
- Is the effective temperature reached by an MSQW of any practical use?

The first two of these will require understanding of fluctuations and moving away from the standard canonical formulation of statistical mechanics. The work presented in this dissertation used a semi-classical analysis to describe the steady-state behaviour of MSQWs, ignoring the role of coherence in the evolution. The last bullet point in the above list is a question about comparing to conventional approaches, which has not yet been studied in the literature. Despite these challenges, it is clear that statistical mechanics presents a practically useful method of discussing the behaviour of strongly interacting spin systems.

8.2.2 Incorporating more information

Given that each anneal is likely to be very short, compared to the time required to remain adiabatic, it is important that each anneal is sufficiently optimised. Cyclic approaches are an attempt to use information

from prior anneals [70, 71, 73, 74]. There are also forward approaches to bias the evolution towards known states [67, 193, 194]. The limitations and successes of these approaches are not well understood – is biasing towards what is likely to be a local optimum a sensible idea? Beyond individual strings, other approaches to incorporate information about the problem comes in the form of variational approaches [40] such as QAOA [14]. In this case previous information is used to calculate observables which are then used to update variational parameters to search for better solutions. As discussed in Chapter 2, these variational approaches are not without their limitations.

In Chapter 4 it was shown how a graph property, specifically the number of triangles in the graph, calculated classically, could be incorporated into a quantum algorithm. This approach did not involve any learning phase. Working out how to incorporate graph properties is no doubt a challenge but a necessary one for CTQO to solve. Consider that superconducting qubits have a fixed topology [195, 196] and Rydberg atoms are typically arranged in a King’s graph [17], and yet there is no prescriptive difference in how the CTQO algorithm should be implemented. This is despite the existence of classical algorithms that cleverly exploit this fixed structure [197, 198]. A reasonable starting point might be how to optimise the performance, based on proposals to embed every Ising problem into three-regular graphs [199] or a planar lattice [190].

8.2.3 Shifting the focus away from ground-state probabilities

Finally, finding the ground state for large optimisation problems is likely to be unachievable and unverifiable. The focus needs to move towards verifiable statements, such as expectation values of local observables. Measuring the quality of solution achieved, as opposed to the small-system scaling of the vanishingly small chance of finding the ground-state, is of more practical interest.

The overall success of CTQO remains undetermined. Too much algorithmic development has focused on adiabaticity, as opposed to what might be realised even on a superficially idealised quantum annealer.

Combinatorial optimisation problems occur frequently and are likely to remain hard to solve. Any advantage or speed-up when it comes to solving them is likely to be beneficial. Quantum algorithms may one day be part of the arsenal of approaches used to tackle these problems. This dissertation has presented a novel approach to understanding current continuous-time approaches, as well as exploring an alternative approach. This presents a positive step forward in understanding the future role of continuous-time quantum algorithms.

Appendix A

Further numerical work on CTQO

A.1 Numerical validation of PSTQA

In Sec. 6.3 it was shown, for a specific Max-cut instance, that the PSTQA equations (Eqs. 6.42-6.45) hold well for that example. Here, we consider other 10-qubit Max-cut instances with different annealing times. The schedule in each case is linear with $A(0) = B(t_f) = 1.3$ and $A(t_f) = B(0) = 0.3$. First, we consider a single 10-qubit example with $t_f = 12$. Focusing on $\langle H_p(t) \rangle$, shown in Fig. A.1, the blue line shows the Schrödinger equation and the pink line the PSTQA equations. There is a gap between the two curves at the beginning of the evolution, as the co-

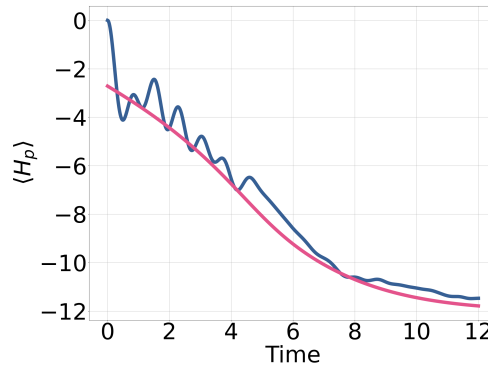


Figure A.1: The time evolution of $\langle H_p(t) \rangle$ for a Max-cut instance on a 10-qubit binomial graph. The schedule is a linear ramp, with a minimum value of 0.3 and maximum value 1.3. The blue line shows the Schrödinger evolution. The pink line shows the solution of the PSTQA equations for this instance.

A. FURTHER NUMERICAL WORK ON CTQO

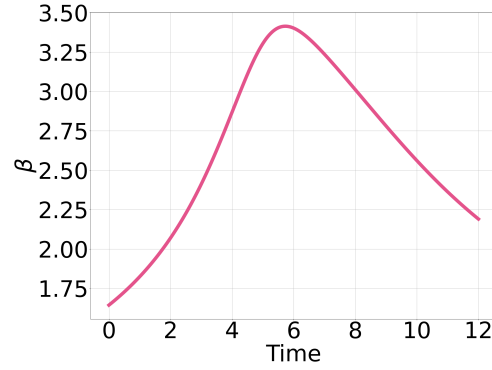


Figure A.2: The inverse temperature for the Max-cut instance considered in Fig. A.1 from the PSTQA equations.

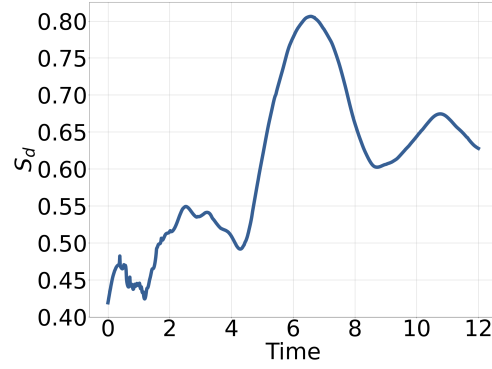


Figure A.3: The diagonal entropy for the Max-cut instance considered in Fig. A.1 from the Schrödinger equations.

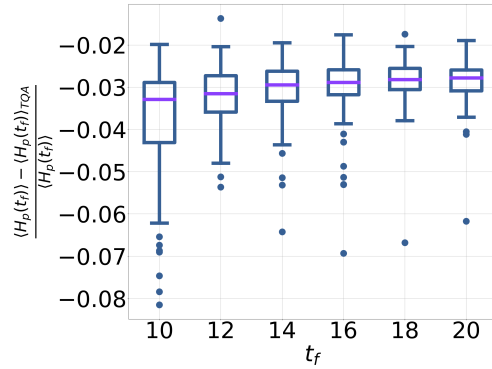


Figure A.4: A box-plot showing $(\langle H_p(t_f) \rangle - \langle H_p(t_f) \rangle_{TQA}) / \langle H_p(t_f) \rangle$ for Max-cut on 10 qubit binomial graphs. The final value of the Schrödinger evolution is denoted by $\langle H_p(t_f) \rangle$. The value predicted by the PSTQA equations is denoted by $\langle H_p(t_f) \rangle_{TQA}$. In each case, a linear schedule is used with $A(0) = B(t_f) = 1.3$ and $A(t_f) = B(0) = 0.3$. At each value of t_f , 90 instances are considered.

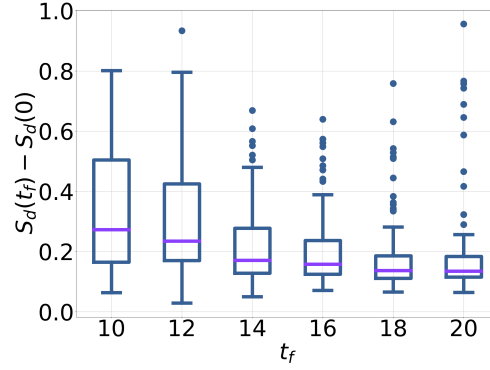


Figure A.5: A box-plot showing the change in diagonal entropy for Max-cut on 10 qubit binomial graphs for the instances shown in Fig. A.4.

herences in the Schrödinger evolution hide the thermal state. At the end of the evolution, the PSTQA equations overestimate the performance of the evolution. The inverse temperature (shown in Fig. A.2) tells a different story to the Max-cut example in Sec. 6.3. In this example, the temperature cools for approximately half the interval before heating up again. The diagonal entropy for this example is shown in Fig. A.3. The evolution is not strictly adiabatic, evidenced by the varying diagonal entropy.

Fig. A.4 shows the error between the error between $\langle H_p(t_f) \rangle$ according to the Schrödinger equation and the prediction from the PSTQA equations ($\langle H_p(t_f) \rangle_{TQA}$) for 90 instances as t_f is changed. Each graph is a 10-qubit binomial graph. The error is typically within a few percent and decreasing on average as the run-time increases. Note that in all instances, the PSTQA equations predict a lower value of $\langle H_p(t_f) \rangle$ than the true value. This is reflected in Fig. A.5 where there is a net increase in diagonal entropy for each instance which is not captured by the PSTQA equations. Despite this, we still see relatively good agreement.

To show that PSTQA is not limited to Max-cut, we consider the Sherrington-Kirkpatrick inspired model (SKM), introduced in Sec. 2.3. For the driver Hamiltonian we take the transverse-field. The schedule under consideration is shown in Fig. A.6. Fig. A.7 compares the value of $\langle H_p(t) \rangle$ according to the Schrödinger equation (blue line) to the PSTQA

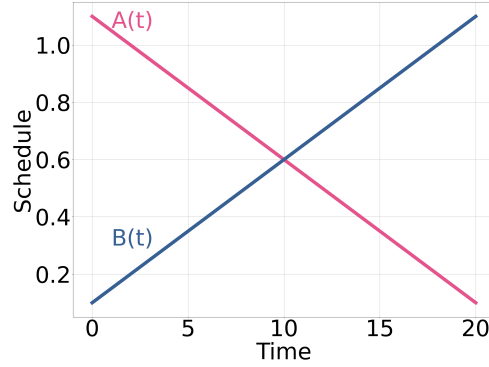


Figure A.6: The schedule used for the 10-qubit SKM instance

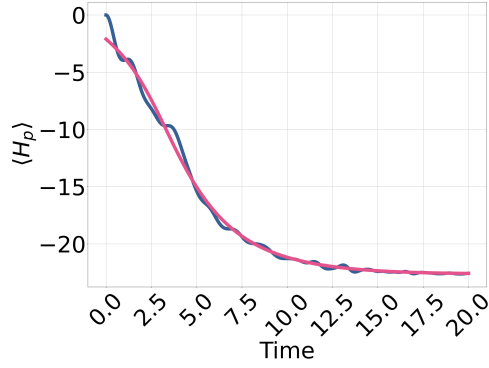


Figure A.7: A comparison of $\langle H_p(t) \rangle$ calculated from the Schrödinger equation (the solid blue line) and the PSTQA equations (the solid pink line) for the 10 qubit SKM instance. The schedule is shown in Fig. A.6

equations (pink line) for a 10 qubit example. There is remarkably good agreement. As in the Max-cut case, Fig. A.8 shows that the diagonal entropy is changing, so the system has not reached the adiabatic limit. Fig. A.9 shows the inverse temperature, which is non-monotonic with time.

Again, we consider 90 instances of the SKM model at various t_f . A similar trend for the error in $\langle H_p(t_f) \rangle$ as the Max-cut instances can be found in Fig. A.10. The change in diagonal entropy can be found in Fig. A.11, which decreases as the run time increases.

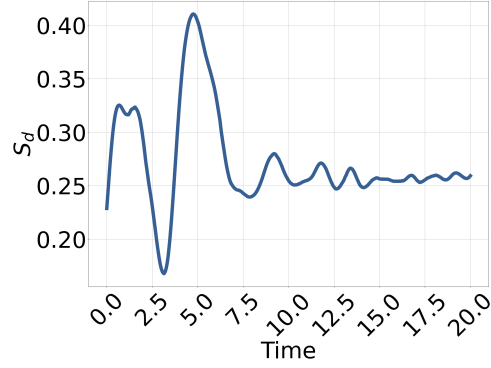


Figure A.8: The diagonal entropy calculated from the Schrödinger equation for the 10 qubit SKM example (Fig. A.7).

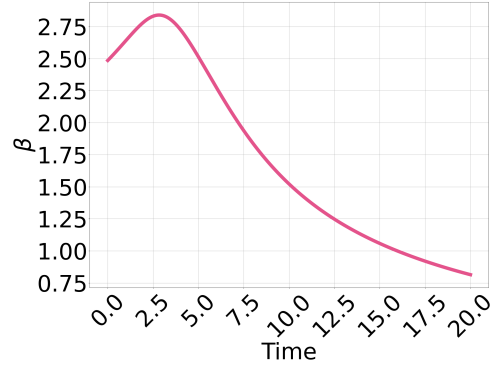


Figure A.9: The inverse temperature calculated from the PSTQA equations for the 10 qubit SKM example (Fig. A.7).

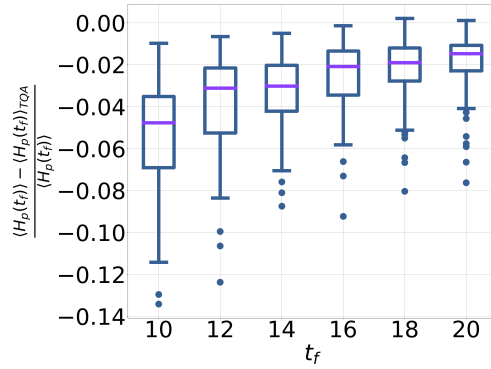


Figure A.10: A box-plot showing $(\langle H_p(t_f) \rangle - \langle H_p(t_f) \rangle_{TQA}) / \langle H_p(t_f) \rangle$ for 10 qubit SKM instances. The final value of the Schrödinger evolution is denoted by $\langle H_p(t_f) \rangle$. The value predicted by the PSTQA equations is denoted by $\langle H_p(t_f) \rangle_{TQA}$. In each case a linear schedule is used with $A(0) = B(t_f) = 1.1$ and $A(t_f) = B(0) = 0.1$. At each value of t_f 90 instances are considered.

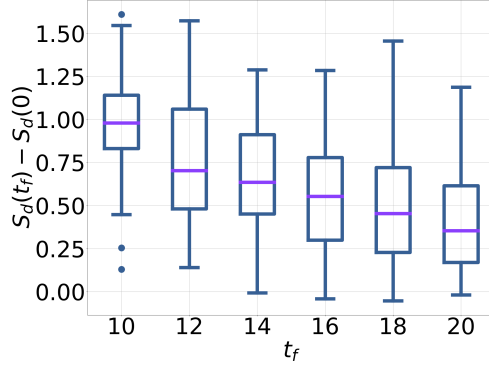


Figure A.11: A box-plot showing the change in diagonal entropy of the 10-qubit SKM instances shown in Fig. A.4.

A.2 Ansatz approaches towards PSTQA

In this section, we show how the PSTQA equations (Eqs. 6.42-6.45) can be tackled through a suitable choice of ansatz. We consider two simple models for the density-of-states. The two models are a Gaussian and an exponentially modified Gaussian. We expect these models to be suitable when the state-vector has significant overlap with energy eigenstates in the middle of the spectrum. These models have no energy cut-off, so become less suitable at low (or high) energies. However, they do allow for some analytic analysis.

A.2.1 The Gaussian model

First, we assume that the density-of-states associated with the Hamiltonian in Eq. 6.40 can be well modelled by a Gaussian distribution:

$$\Omega(\varepsilon, t) d\varepsilon = \frac{1}{\sqrt{2\pi\sigma^2(t)}} e^{-\frac{(\varepsilon - \mu(t))^2}{2\sigma^2(t)}} d\varepsilon, \quad (\text{A.1})$$

where $\mu(t)$ is the mean and $\sigma^2(t)$ the variance of the density-of-states. A Gaussian density-of-states has been observed to be a good approximation for the density-of-states in the non-integrable setting for a number of models [137, 138, 139]. The moments of the density-of-states can be calculated directly from the eigenvalues of $H_{TQA}(t)$, denoted be E_k .

Again, let D be the dimension of the state space. And $\text{Tr}' = 1/D \text{Tr}$ be the normalised Trace, with the scaled operators $\tilde{H}_p = H_p - \text{Tr}' H_p$ and $\tilde{H}_d = H_d - \text{Tr}' H_d$. The mean and variance of the density-of-states are then

$$\begin{aligned}\mu(t) &= \frac{1}{D} \sum_k E_k(t) \\ &= \text{Tr}' H_{TQA}(t) \\ &= A(t) \text{Tr}' H_d + B(t) \text{Tr}' H_p\end{aligned}\tag{A.2}$$

$$\begin{aligned}\sigma^2(t) &= \frac{1}{D} \sum_k E_k^2(t) \\ &= \frac{1}{D} \text{Tr} H_{TQA}^2(t) - \mu(t)^2 \\ &= A^2(t) \text{Tr}' \tilde{H}_d^2 + B^2(t) \text{Tr}' \tilde{H}_p^2 \\ &\quad + 2A(t)B(t) \text{Tr}' \tilde{H}_d \tilde{H}_p.\end{aligned}\tag{A.3}$$

Evaluating the partition function gives:

$$\begin{aligned}\mathcal{Z}(t) &= \int_{-\infty}^{\infty} e^{-\beta(t)\varepsilon} \Omega(\varepsilon, t) d\varepsilon \\ &= \frac{1}{\sqrt{2\pi\sigma^2(t)}} \int_{-\infty}^{\infty} e^{-\beta(t)\varepsilon} e^{-\frac{(\varepsilon - \mu(t))^2}{2\sigma^2(t)}} d\varepsilon \\ &= e^{-\beta(t)\mu(t) + \frac{\beta(t)^2\sigma^2(t)}{2}}.\end{aligned}\tag{A.4}$$

Evaluating $\langle H_{TQA}(t) \rangle$ gives:

$$\begin{aligned}\langle H_{TQA}(t) \rangle &= -\frac{\partial \ln \mathcal{Z}(t)}{\partial \beta} \\ &= \mu(t) - \beta\sigma^2(t)\end{aligned}\tag{A.5}$$

Hence:

$$\beta(t) = \frac{\mu(t) - \langle H_{TQA}(t) \rangle}{\sigma^2}.\tag{A.6}$$

Evaluating $\langle H_d(t) \rangle$ and $\langle H_p(t) \rangle$ gives:

$$\begin{aligned} \langle H_d(t) \rangle &= -\frac{1}{\beta(t)} \frac{\partial \ln \mathcal{Z}(t)}{\partial A} \\ &= \text{Tr}' H_d - \beta(t) \left(A(t) \text{Tr}' \tilde{H}_d^2 + B(t) \text{Tr}' \tilde{H}_d \tilde{H}_p \right) \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \langle H_p(t) \rangle &= -\frac{1}{\beta(t)} \frac{\partial \ln \mathcal{Z}(t)}{\partial B} \\ &= \text{Tr}' H_p - \beta(t) \left(B(t) \text{Tr}' \tilde{H}_p^2 + A(t) \text{Tr}' \tilde{H}_d \tilde{H}_p \right). \end{aligned} \quad (\text{A.8})$$

Substituting $\beta(t)$, $\langle H_d(t) \rangle$ and $\langle H_p(t) \rangle$ into Eq. 6.42 gives:

$$\begin{aligned} \frac{d\langle H_{TQA}(t) \rangle}{dt} &= \dot{A}(t) \langle H_d(t) \rangle + \dot{B}(t) \langle H_p(t) \rangle \\ &= \frac{d\mu}{dt} - \frac{\mu - \langle H_{TQA}(t) \rangle}{2\sigma^2} \frac{d\sigma^2}{dt} \end{aligned} \quad (\text{A.9})$$

Integrating the above expression gives:

$$\langle H_{TQA}(t) \rangle = \mu(t) + c\sigma(t), \quad (\text{A.10})$$

where c is the constant of integration, which can be fixed using the boundary condition $\langle H_{TQA}(0) \rangle = A(0) \langle \psi_i | H_d | \psi_i \rangle + B(0) \langle \psi_i | H_p | \psi_i \rangle$, where $|\psi_i\rangle$ is the initial state. Therefore:

$$c = \frac{\langle H_{TQA}(0) \rangle - \mu(0)}{\sigma(0)}. \quad (\text{A.11})$$

With an expression for $\langle H_{TQA}(t) \rangle$, evaluating $\beta(t)$ and $\langle H_p(t) \rangle$ becomes trivial:

$$\beta(t) = \frac{-c}{\sigma(t)} \quad (\text{A.12})$$

and

$$\langle H_p \rangle = \text{Tr}' H_p + \frac{c}{\sigma(t)} \left(B(t) \text{Tr}' \tilde{H}_p^2 + A(t) \text{Tr}' \tilde{H}_d \tilde{H}_p \right) \quad (\text{A.13})$$

A.2.2 The Gaussian model applied to Max-cut

The simple structure of the Max-cut problem and the transverse field allow for further simplification. Evaluating the moments gives:

$$\begin{aligned}\mu(t) &= 0 \\ \sigma^2(t) &= A^2(t)n + B^2(t)\kappa_2,\end{aligned}$$

where κ_2 is the number of edges in the Max-cut graph and n is the number of nodes. Hence:

$$\begin{aligned}\beta &= \frac{nA(0)}{\sqrt{A^2(0)n + B^2(0)\kappa_2}\sqrt{A^2(t)n + B^2(t)\kappa_2}} \\ \langle H_{TQA}(t) \rangle &= \frac{-nA(0)\sqrt{A^2(t)n + B^2(t)\kappa_2}}{\sqrt{A^2(0)n + B^2(0)\kappa_2}} \\ \langle H_p \rangle &= \frac{-nA(0)B(t)\kappa_2}{\sqrt{A^2(0)n + B^2(0)\kappa_2}\sqrt{A^2(t)n + B^2(t)\kappa_2}}.\end{aligned}$$

The Gaussian case is primarily of interest as it can be handled analytically. It does not take into account any frustration in the system. In the next section, we explore a model that begins to take this into account.

A.2.3 The exponentially modified Gaussian model

The above can be repeated for an exponentially modified Gaussian density-of-states. This model incorporates skewness into the density-of-states model, but is less tractable. The partition function for this model is given by:

$$\mathcal{Z}(t) = \left(1 + \frac{\beta(t)}{\lambda(t)}\right)^{-1} e^{-\nu(t)\beta(t) + \frac{1}{2}\beta(t)^2 s^2}, \quad (\text{A.14})$$

where $\nu(t)$, $s(t)$ and $\lambda(t)$ are fitting parameters related to the mean ($\mu(t)$), variance ($\sigma^2(t)$) and skewness $\gamma(t)$ of the distribution:

$$\nu(t) = \mu(t) - \sigma(t) \left(\frac{\gamma(t)}{2} \right)^{\frac{1}{3}} = \mu(t) - \Delta(t), \quad (\text{A.15})$$

$$s^2(t) = \sigma^2 \left(1 - \left(\frac{\gamma(t)}{2} \right)^{\frac{2}{3}} \right) = \sigma^2(t) - \Delta^2(t), \quad (\text{A.16})$$

$$\lambda(t) = \frac{1}{\sigma(t)} \left(\frac{\gamma(t)}{2} \right)^{-\frac{1}{3}} = \frac{1}{\Delta(t)}, \quad (\text{A.17})$$

where:

$$\mu(t) = \text{Tr}' H_{TQA} \quad (\text{A.18})$$

$$\sigma^2(t) = \text{Tr}' (H_{TQA} - \text{Tr}' H_{TQA})^2, \quad (\text{A.19})$$

$$\Delta(t) = \frac{1}{2} \left(\text{Tr}' (H_{TQA} - \text{Tr}' H_{TQA})^3 \right)^{1/3}. \quad (\text{A.20})$$

From the partition function, it is straight forward to estimate the relevant expectation values:

$$\langle H_{TQA}(t) \rangle = \frac{\beta^2 \Delta^3(t)}{1 + \beta(t) \Delta(t)} + \mu(t) - \beta(t) \sigma^2(t). \quad (\text{A.21})$$

Inverting this equation gives:

$$\beta(t) = \frac{-\sigma^2 - \Delta (\langle H_{TQA}(t) \rangle - \mu) + \omega}{2\Delta(t) (\sigma^2(t) - \Delta^2(t))}, \quad (\text{A.22})$$

where

$$\begin{aligned} \omega = & \left\{ \left[\sigma^2(t) + \Delta(t) (\langle H_{TQA}(t) \rangle - \mu(t)) \right]^2 \right. \\ & \left. + 4\Delta (\langle H_{TQA}(t) \rangle - \mu(t)) (\Delta^2(t) - \sigma^2(t)) \right\}^{1/2}. \end{aligned} \quad (\text{A.23})$$

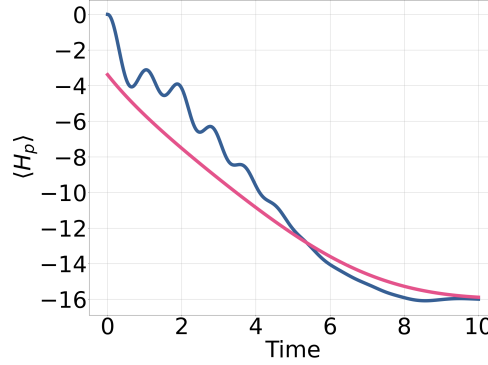


Figure A.12: The evolution of $\langle H_p(t) \rangle$ for a Max-cut instance on 13 qubit binomial graph. The schedule is linear with $A(0) = B(t_f = 10) = 1.1$ and $A(t_f = 10) = B(0) = 0.1$. The blue line shows the Schrödinger evolution, and the pink line the prediction using an exponentially modified Gaussian ansatz for the density-of-states.

Calculating $\langle H_p \rangle$ gives:

$$\begin{aligned} \langle H_p(t) \rangle = \partial_B \mu - \frac{\Delta(t)\beta(t)}{1 + \Delta\beta(t)} \partial_B \Delta \\ - \beta(t) (\sigma(t) \partial_B \sigma(t) - \Delta(t) \partial_B \Delta(t)). \end{aligned} \quad (\text{A.24})$$

The expression for $\langle H_d(t) \rangle$ is the same but with B swapped for A . Combining the above expressions, the energy of the system (Eq. 6.46) can be found using numerical integration, without resorting to full numerical integration of the state-vector.

A.2.4 The exponentially modified Gaussian model applied to Max-cut

For Max-cut with a transverse field (and no terms proportional to the identity):

$$\mu = 0 \quad (\text{A.25})$$

$$\sigma^2(t) = A^2(t)n + B^2(t)\kappa_2 \quad (\text{A.26})$$

$$\Delta^3(t) = (3\kappa_3)B^3(t). \quad (\text{A.27})$$

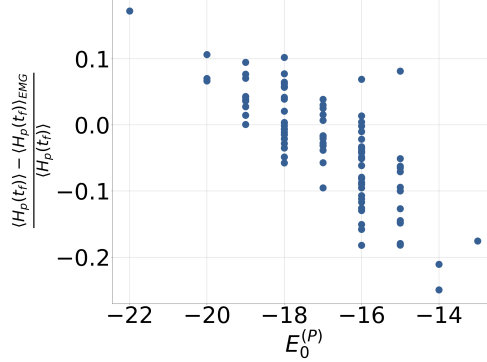


Figure A.13: The error between $\langle H_p(t_f) \rangle$ from the Schrödinger equation and the exponentially modified Gaussian model, $\langle H_p(t_f) \rangle_{EMG}$. The figure shows 100 Max-cut instances on 13 qubit binomial graphs. The schedule is linear with $A(0) = B(t_f = 10) = 1.1$ and $A(t_f = 10) = B(0) = 0.1$. The prediction from the exponentially modified Gaussian is denoted by $\langle H_p(t_f) \rangle_{EMG}$, and the true value by $\langle H_p(t_f) \rangle$. The x-axis shows $E_{(0)}^p$, the ground state energy of H_p .

Here κ_2 is the number of edges and κ_3 is the number of triangles in the Max-cut graph. Fig. A.12 shows the Schrödinger evolution (blue line) and the prediction from using an exponentially modified Gaussian ansatz (pink line). The instance considered is a 13-qubit binomial graph with a linear schedule with $A(0) = B(t_f = 10) = 1.1$ and $A(t_f = 10) = B(0) = 0.1$. There is good agreement, especially at the end of the evolution. Fig. A.13 shows the error between the true value of $\langle H_p(t_f) \rangle$ and the prediction from the ansatz $\langle H_p(t_f) \rangle_{EMG}$ for 100 13 qubit instances. The schedule is the same as before. The x-axis shows the ground state energy of the problem Hamiltonian, which is correlated with the error.

A.3 Warm-starting CTQWs

In order to provide numerical evidence of the challenges of warm-starting CTQWs, again we focus on Max-cut and the SKM. For both cases, we take H_d to be the transverse-field driver.

In this section, we focus on 12-qubit examples. Fig. A.14a shows a Max-cut instance and Fig. A.14b an SKM instance of a warm-started CTQW.

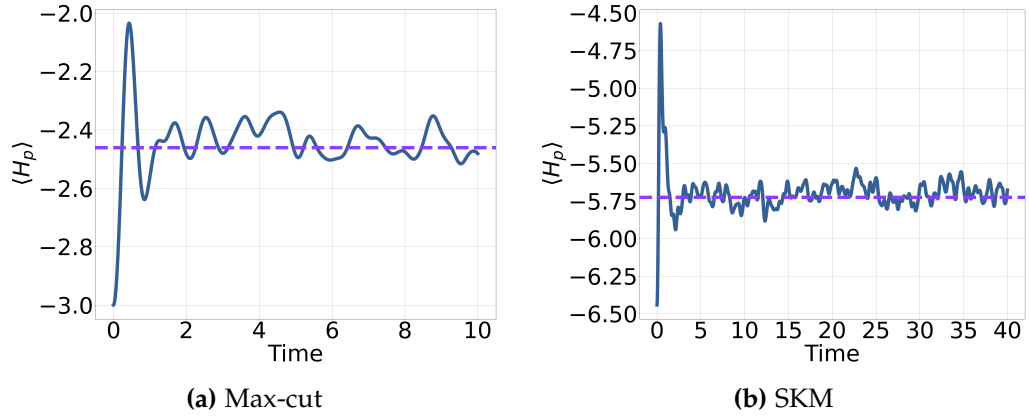


Figure A.14: The time-evolution of $\langle H_p \rangle$ for a 12 qubit instance. The dashed purple line shows the infinite time average, $\overline{\langle H_p \rangle}$.

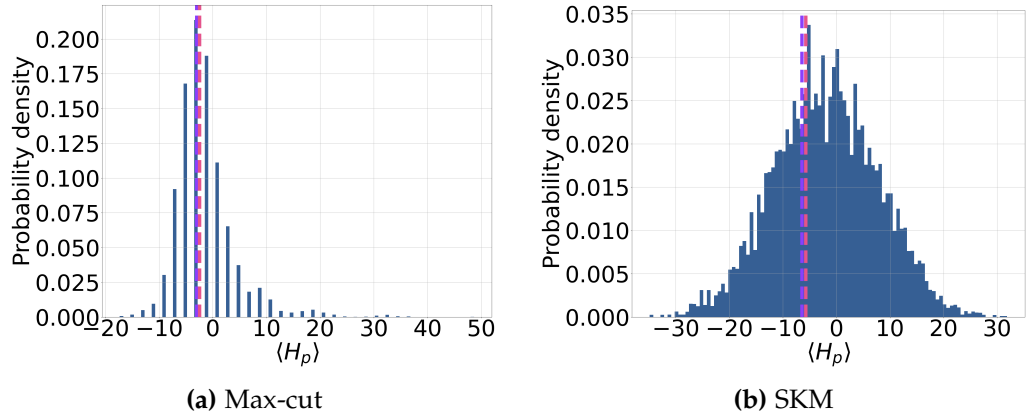


Figure A.15: The time-averaged distribution of $\langle H_p \rangle$ for the problem instances shown in Fig. A.14. The dashed purple line shows the original value of $\langle H_p \rangle$. The dashed pink line shows the average of the distribution.

In both cases $g = 1$, where g is the coefficient in front of H_d . The evolution after some time approaches a steady state, as expected for a time-independent Hamiltonian. The dashed purple line shows the infinite time average of $\langle H_p \rangle$, denoted by $\overline{\langle H_p \rangle}$. As predicted by Assumption 4, $\langle H_p \rangle$ is always greater than its initial value.

Since the initial state is an energy eigenstate, the uncertainty can only increase, giving the warm-started CTQW a chance to find better solutions. To see this for the warm-started CTQWs, we look at the probability distribution of observing a given value of $\langle H_p \rangle$ for the infinite time aver-

A. FURTHER NUMERICAL WORK ON CTQO

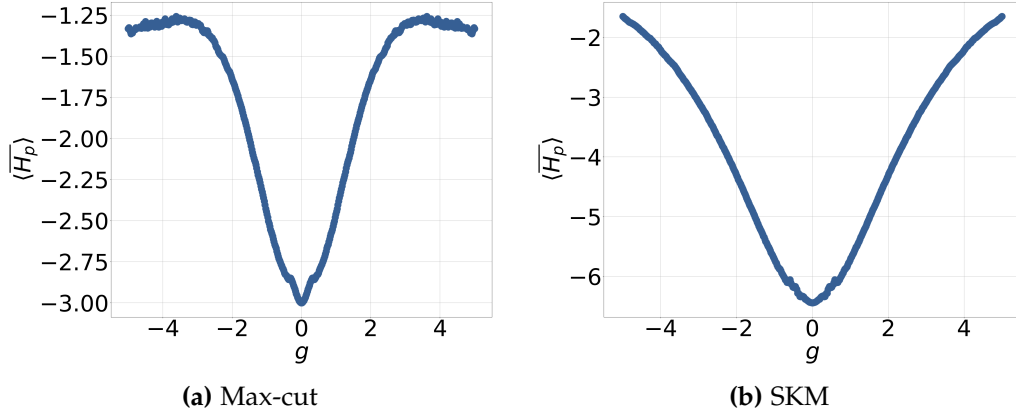


Figure A.16: The time-averaged value of $\langle H_p \rangle$ as g is varied for the Max-cut instance considered in Fig. A.14.

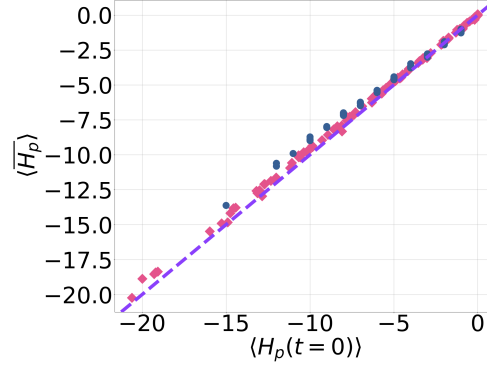


Figure A.17: The initial value of $\langle H_p \rangle$ for a warm-started CTQW compared to the infinite time averaged value. For each instance $g = 0.5$. The blue circles show 100 Max-cut instances. The pink diamonds show 100 SKM instances. The number of qubits is 12 in all cases. The initial state is randomly selected given that it satisfies Eq. 6.53. The dashed purple line is the “ $y = x$ ” line for ease of visual comparison.

aged density operator. We consider the same examples as Fig. A.14. The distribution of $\langle H_p \rangle$ for the Max-cut instance can be found in Fig. A.15a and the SKM instance in Fig. A.15b. In both cases, despite the average value of $\langle H_p \rangle$ having increased, there is significant overlap with states with a lower value of $\langle H_p \rangle$ than the initial state.

To demonstrate that the average increase in $\langle H_p \rangle$ is not specific to the choice of g for these instances, Fig. A.16a and Fig. A.16b show how $\langle H_p \rangle$ varies with g . The instances are the same as Fig. A.14. Both demonstrate clear heating, as for any value of g the value of $\langle H_p \rangle$ increases or stays

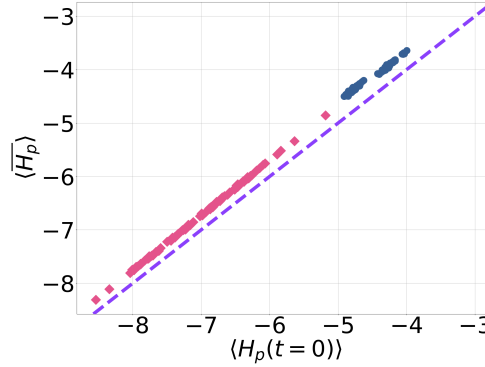


Figure A.18: The initial value of $\langle H_p \rangle$ for a warm-started CTQW compared to the infinite time averaged value. For each instance $g = 0.5$. The blue circles show 100 Max-cut instances. The pink diamonds show 100 SKM instances. The number of qubits is 12 in all cases. The initial state is given by Eq. A.28.

the same compared to its initial value. The initial value of $\overline{\langle H_p \rangle}$ corresponds to $g = 0$. So for the cyclic process set out in Sec. 6.4.1 there is an increase in energy, this corresponds to heating.

Fig. A.17 shows how $\overline{\langle H_p \rangle}$ compares with the initial value of $\langle H_p(t=0) \rangle$ for 100 instances of the Max-cut problem and 100 instances of the SKM problem. In each instance $g = 0.5$. The dashed purple line marks no change in $\langle H_p \rangle$. For a few initial states, there is a minor cooling effect from the CTQW. This is perhaps reflective of the problem having more structure than the bell shape sketched in Fig. 3.1.

In each instance so far we have shown what has happened when the system is initialised randomly in an eigenstate of H_p , that satisfies the condition $\langle s | H_p | s \rangle < 0$ where $|s\rangle$ is the initial-state. For a few instances in Fig. A.17 we have seen $\overline{\langle H_p \rangle}$ decrease — does this imply cooling and contradict the work in the main section? Firstly, we have taken Planck’s principle to be a physically motivated principle and broadly true. Secondly, this cooling effect disappears once the system is averaged over the full ensemble, taking into account all possible starting states. The

full ensemble can be taken into account by using the density operator:

$$\rho_0 = \frac{1}{\mathcal{N}} \sum_{s: \langle s | H_p | s \rangle < \text{Tr}' H_p} |s\rangle \langle s|, \quad (\text{A.28})$$

where $|s\rangle$ is an eigenstate of H_p with eigenvalue s . The normalisation \mathcal{N} of the initial state is

$$\mathcal{N} = \sum_{s: \langle s | H_p | s \rangle < 0} 1. \quad (\text{A.29})$$

Repeating Fig. A.17 with Eq. A.28 gives Fig. A.18. Note that ρ_0 is a passive state. From Fig. A.18 it is clear that by averaging over the ensemble given by Eq. A.28 there is no exception.

A.4 Numerically observing cyclic cooling

Sec. 6.4 made two predictions:

1. Cyclic processes lead to heating, so RQA should lead to a greater value of $\langle H_p \rangle$.
2. A cyclic process might achieve cooling of $\langle H_p \rangle$ with the introduction of a third term in the Hamiltonian.

In this section, we numerically investigate this for the Max-cut problem and the SKM.

To start, we consider the first prediction, that RQA without any biasing term leads to heating. Before investigating heating, we outline our model of RQA in more detail. The initial state is the ensemble

$$\rho_0 = \sum_s p(s) |s\rangle \langle s|, \quad (\text{A.30})$$

where $|s\rangle$ is an eigenstate of H_p with eigenvalue s . Denoting the unitary associated with one RQA cycle to be U_{cyc} , then the transition probability

between states $|s\rangle$ and $|j\rangle$ is given by

$$p(j|s) = |\langle j| U_{\text{cyc}} |s\rangle|^2. \quad (\text{A.31})$$

After each stage, there is some selection criterion to determine if a string is kept or if the RQA cycle is repeated with the initial string. In this work, we only keep states that lower the energy.

The state after k iterations and measurement but before post-selection is given by

$$\rho_k = \frac{1}{\mathcal{N}_k} \sum_{s_k} \sum_{s_{k-1} < \dots < s_1 < s_0} p(s_k|s_{k-1}) p(s_{k-1}|s_{k-2}) \dots p(s_1|s_0) p(s_0) |s_k\rangle \langle s_k|, \quad (\text{A.32})$$

where \mathcal{N}_k normalises the state. The probability of finding a lower state from this ensemble is

$$p_{\text{suc}}(k) = \frac{1}{\mathcal{N}_k} \sum_{s_k < s_{k-1} < \dots < s_1 < s_0} p(s_k|s_{k-1}) p(s_{k-1}|s_{k-2}) \dots p(s_1|s_0) p(s_0). \quad (\text{A.33})$$

This sets a rough estimate for the number of iterations to lower the energy. If $1/p_{\text{suc}}(k)$ becomes greater than some cut-off, we terminate the RQA process. If the approach is not terminated, the state that is fed into the next cyclic iteration is,

$$\rho_k^{(ps)} = \frac{1}{\mathcal{N}'_k} \sum_{s_k < s_{k-1} < \dots < s_1 < s_0} p(s_k|s_{k-1}) p(s_{k-1}|s_{k-2}) \dots p(s_1|s_0) p(s_0) |s_k\rangle \langle s_k|. \quad (\text{A.34})$$

This completes our model of RQA.

To illustrate the above analysis, we consider a single 10-vertex Max-cut instance, take $p(s)$ to be a uniform distribution on the interval $s <$

A. FURTHER NUMERICAL WORK ON CTQO

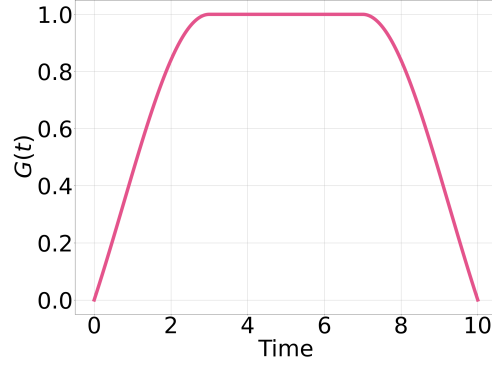


Figure A.19: The schedule $G(t)$ appended to the driver Hamiltonian, used for the cyclic processes described in Appendix A.4.

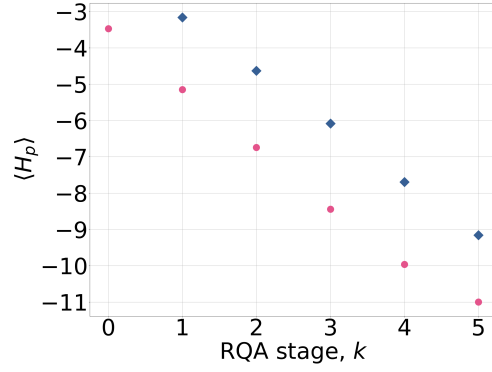


Figure A.20: RQA for a 10 node Max-cut graph. The pink dots show $\langle H_p \rangle$ for the post-selected distribution. The blue diamonds show $\langle H_p \rangle$ as sampled from application of U_{cyc} .

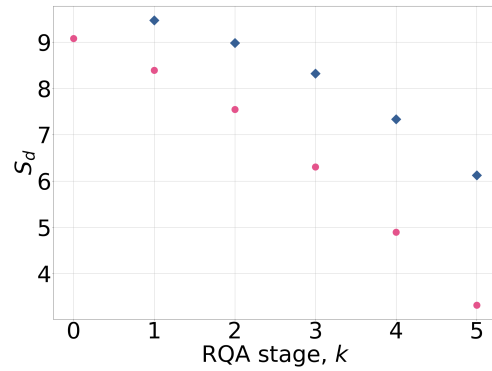


Figure A.21: The diagonal entropy S_d for RQA applied to a 10-node Max-cut graph. The pink dots show S_d for the post-selected distribution. The blue diamonds show S_d as sampled after application of U_{cyc} . The logarithm used to calculate the diagonal entropy is taken to be base 2.

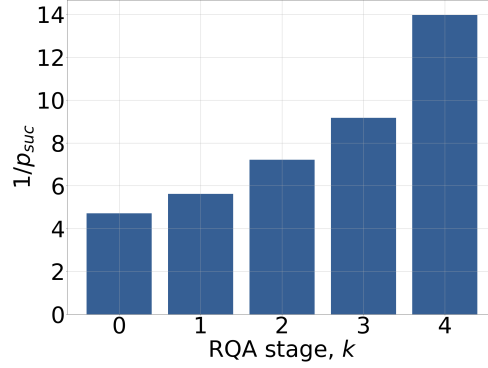


Figure A.22: The inverse probability (Eq. A.33) for each stage of RQA on a 10-vertex Max-cut graph.

$\text{Tr}' H_p$. The driver Hamiltonian is taken to be H_{TF} . The schedule for the drive $G(t)$ is taken to be a square Gaussian, shown in Fig. A.19. For this instance $\langle H_p \rangle$ can be seen in Fig. A.20. The pink circles show $\langle H_p \rangle$ for the state after post-selection. The blue diamonds show $\langle H_p \rangle$ sampled after each single application of U_{cyc} . At each stage $\langle H_p \rangle$ after the cyclic quantum process is greater than the post-selected state, therefore the cyclic quantum process is producing on average worse quality states. This is numeric evidence of heating at each stage. For this instance, RQA reaches the ground-state.

The diagonal entropy for the same process is shown in Fig. A.21. Again, the blue diamonds show the diagonal entropy sampling directly after an application of U_{cyc} . The pink dots show the diagonal entropy of the post selected state. As is clear at each RQA leads to an increase of diagonal entropy corresponding to a broadening of the distribution.

Finally, Fig. A.22 shows, $1/p_{\text{suc}}$ which gives a sense of the average number of shots required at each stage. The randomness U_{cyc} introduces allows RQA to find the ground state with fewer classical evaluations of H_p . This is in spite of the fact, that U_{cyc} results in a larger value of $\langle H_p \rangle$, as we have argued.

Having explored RQA without a bias, we now introduce biased quantum annealing (BQA). Simulating the full ensemble at each stage is com-

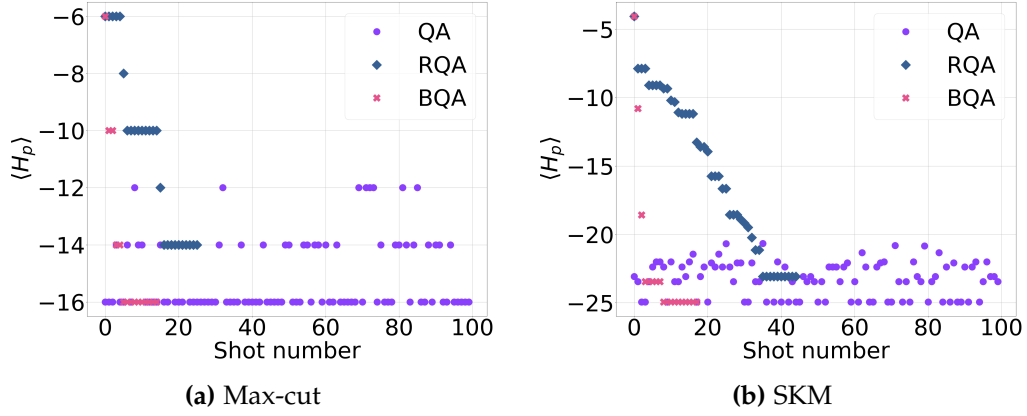


Figure A.23: The performance of QA (purple circles), RQA (blue diamonds) and BQA (pink crosses) for a 12 qubit example.

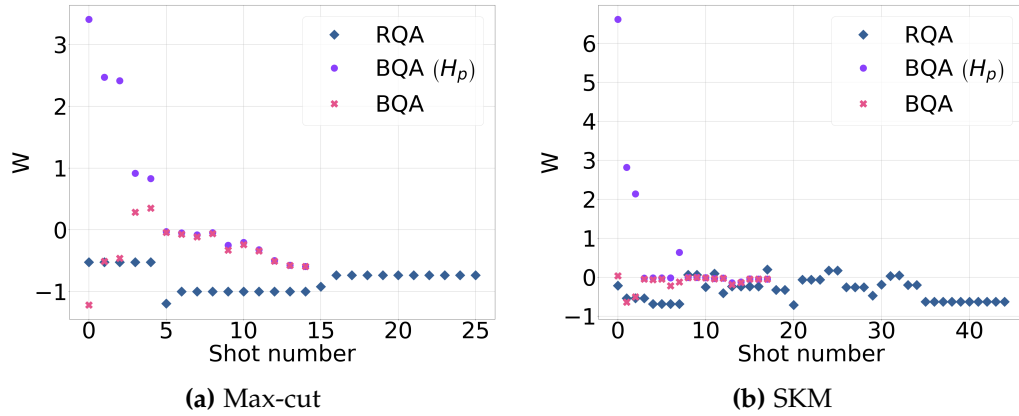


Figure A.24: The extractable work, W , for the 12-qubit instances considered in Fig. A.23. The blue diamonds correspond to RQA. The pink crosses BQA. The purple circles show W for BQA neglecting the change in energy of $H_b^{(l)}$, i.e, the change in $\langle H_p \rangle$.

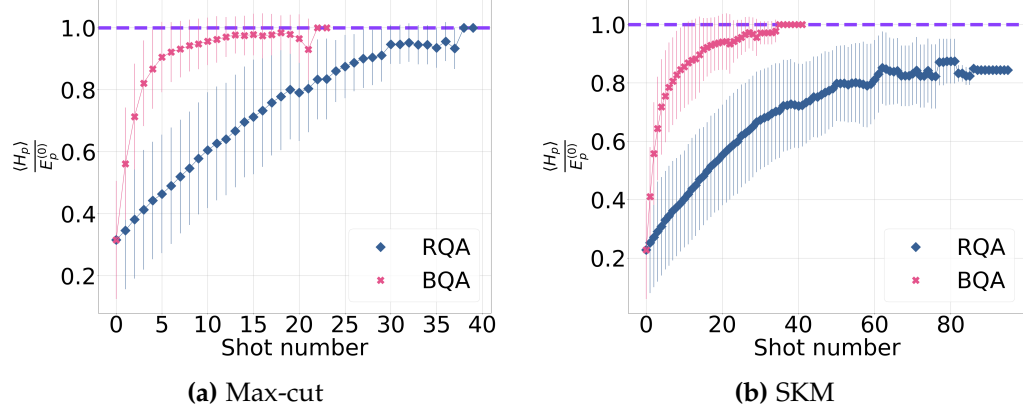


Figure A.25: The average approximation ratio for the 12 qubit instances at each shot number. Since the approaches terminate at different shot numbers, the number of instances decreases. The pink crosses show the average approximation ratio for BQA. The blue diamonds show the average approximation ratio for RQA. The lines show one standard deviation. The dashed purple line shows an approximation ratio of 1. The final decrease in approximation ratio in BQA reflects more difficult problems for BQA, since it has not terminated after many shots.

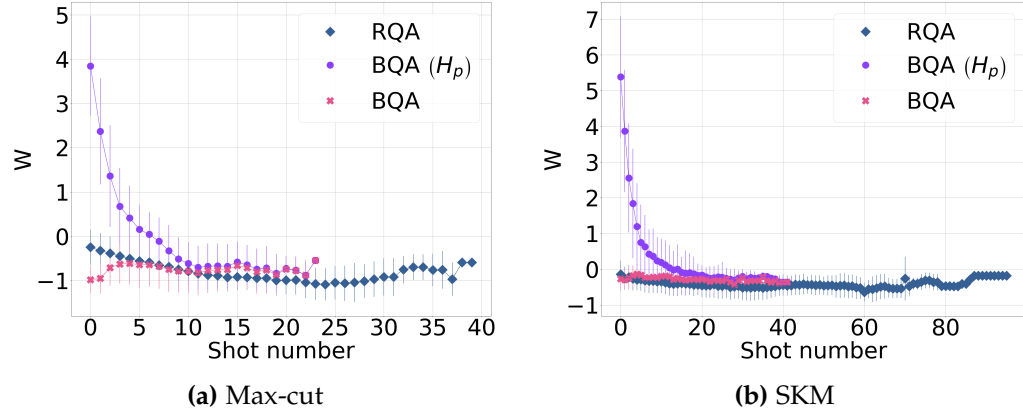


Figure A.26: The extractable work for the 12 qubit instances at each shot number. Since the approaches terminate at different shot numbers, the number of instances decreases. The pink crosses show the average extractable work for the BQA protocol. The blue diamonds show the average extractable work for RQA. The lines show one standard deviation. The purple circles show the change in $\langle H_p \rangle$ for BQA.

putationally more expensive than a single state-vector, so from here on the simulations use sampling as if the approach was being evaluated on an actual quantum device. The problems considered consist of 12 qubits. For each instance, up to k_{\max} shots are taken. The initial state is chosen by randomly selecting states until condition Eq. 6.53 is met. If the result of a run produces a better quality solution than the initial state, the initial state is updated to be this state. If the initial state is not updated after k runs, the algorithm is assumed to have converged. For the numerics, we take $k_{\max} = 100$ and $k = 10$. The schedule for the drive $G(t)$ is taken to be a square Gaussian, shown in Fig. A.19.

For BQA the biasing term is taken to be $H_b^{(l)}$ (i.e. Eq. 6.71). The initial value of α is taken to be $\alpha_0 = \sqrt{\text{Tr}'(H_p^2)}$. This is assumed to be typically an overshoot, so α is decreased with each shot if $\langle H_p \rangle$ does not decrease. We take α to decrease linearly by α_0/k each time.

With an annealing time of $t_a = 10$ the problems are typically easy for QA. We consider 100 instances of Max-cut and SK. QA found the ground state with 100 shots or fewer for 99 of the Max-cut instances and all the SKM instances. RQA found the ground state for 22 of the Max-cut instances and 12 of the SKM instances. BQA found the ground state for 78 of the Max-cut instances and 48 of the SKM instances.

Fig. A.23 shows a specific Max-cut and SKM instance. The purple circles show the result from QA. Though clearly not adiabatic, it finds the ground state in both cases. The blue diamonds show the initial value of $\langle H_p \rangle$ for RQA. In both cases, the termination condition is reached before the algorithm finds the ground state. The pink crosses show BQA. BQA managed to find the ground state in both cases. BQA rapidly converges towards the ground state compared to RQA and terminates within 20 shots.

The above shows that BQA and RQA can tackle combinatorial optimisation problems. The aim of this section is to discuss heating. Fig. A.24 shows the extractable work for each shot for the instances considered

in Fig. A.23. The RQA examples generally show a negative value of extractable work, meaning that $\langle H_p \rangle$ decreases — the blue diamonds in Fig. A.24. BQA also shows heating (i.e, $W < 0$), the pink crosses in Fig. A.24. There are some instances of cooling in the Max-cut instance for BQA for two shots. The purple circles show the change in $\langle H_p \rangle$ for the BQA protocol. We see $\langle H_p \rangle$ is decreasing as a result of BQA, i.e. cooling of $\langle H_p \rangle$.

Finally, we consider the numerics for all the 100 Max-cut instances and 100 SKM instances. Fig. A.25 shows the approximation ration averaged over all instances. The approximation ratio is defined as $\langle H_p \rangle$ divided by the ground state energy of H_p . If the approach finds the ground state, the approximation ratio is 1 (and cannot exceed 1). In both cases, BQA converges much faster than RQA.

Fig. A.26 shows the extractable work averaged over all the instances. We observe heating for both protocols as predicted by Assumption 4. However, at least for the first few shots, BQA is able to achieve cooling of $\langle H_p \rangle$.

We have numerically demonstrated, with this set-up, that RQA leads to heating. This does not rule out RQA finding better solutions, as evidenced by Fig. A.25. It does suggest there is scope for improvement and cooling $\langle H_p \rangle$ by the addition of a third term.

Appendix B

Further numerical work on rapid Hamiltonians for optimisation

This chapter provides further numerical evidence of the optimal state-transfer inspired approaches outlined in Chapter 7. This appendix focuses on Sherrington-Kirkpatrick inspired models (SKMs), introduced in Sec. 2.3, as well as further discussion of where QAOA $p = 1$ outperformed H_1 .

B.1 A further numerical study on a Sherrington-Kirkpatrick inspired model

To provide further evidence of the applicability of the approaches introduced in Chapter 7, in this appendix we repeat the numerical experiments performed on Max-cut for SKMs.

Starting with assessing the performance of H_1 , Fig. B.1 shows the performance of H_1 on 100 randomly generated instances of the SKM. The approximation ratio appears to have little dependence on the problem size for more than 7 qubits, an indicator that the dynamics under H_1 is approximately local for these times. The ground-state probability also appears to decline exponentially (Fig. B.1b). The optimal times can be found in Fig. B.1c. It appears that the optimal time tends to a constant value (or a small range of values), with $T < 1$.

B. FURTHER NUMERICAL WORK ON RAPID HAMILTONIANS FOR OPTIMISATION

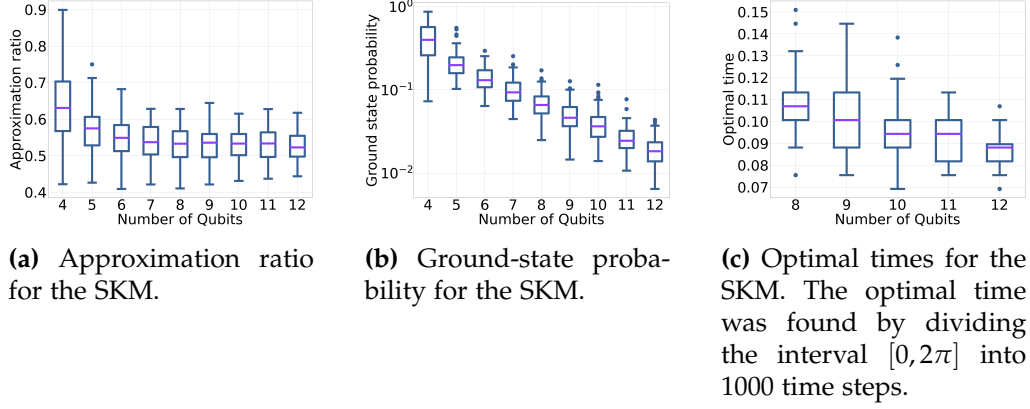


Figure B.1: Performance of H_1 on 100 randomly-generated instances of SKM.

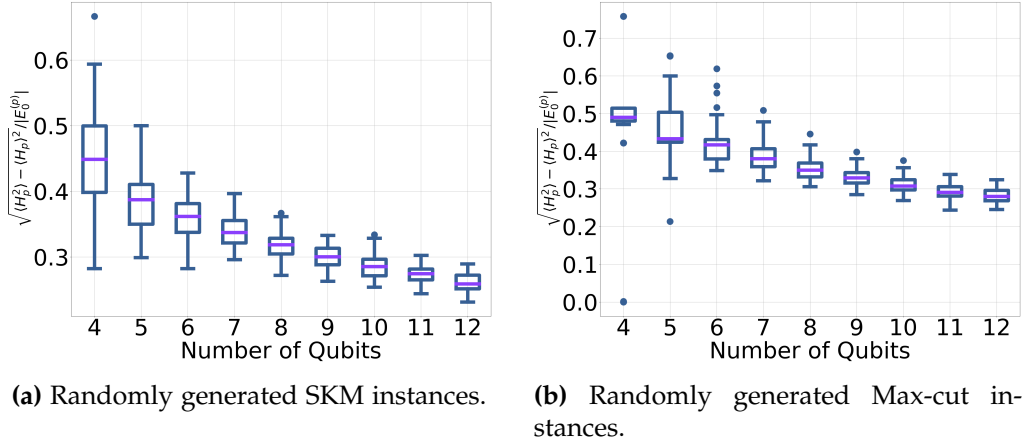
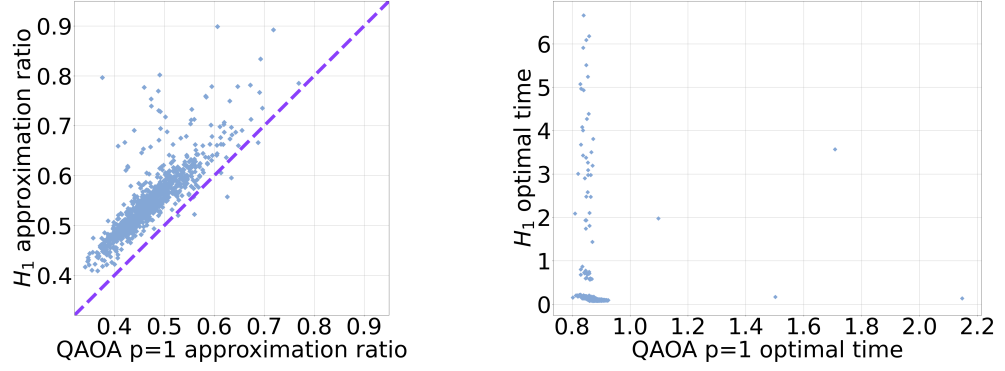


Figure B.2: Width of the final distribution, σ for randomly generated instances of Max-cut and SKM.

Knowing the width of the distribution associated with the approximation ratio can also be useful. This is shown for SKM as well as the Max-cut instances on randomly generated graphs in Fig. B.2. The width, σ is non-zero, suggesting the final state is not a computational basis state.

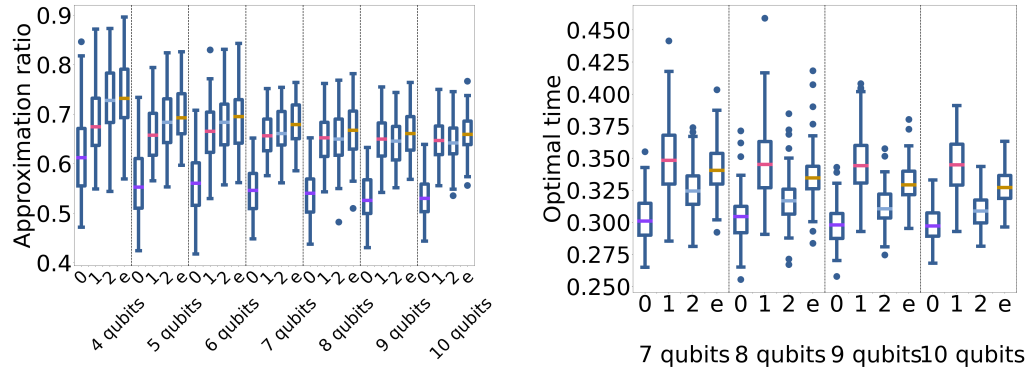
In Fig. B.3 the performance of H_1 is directly compared to QAOA $p = 1$ on 100 instances. For a handful of problems with the SKM, QAOA $p = 1$ outperformed H_1 , but for the vast majority of problem instances H_1 performed better for both approximation ratio and optimal time. In Appendix B.2 we elaborate further on the exceptions.

B.1. A further numerical study on a Sherrington-Kirkpatrick inspired model



(a) Approximation ratio comparison for the SKM. (b) Optimal time comparison for the SKM.

Figure B.3: Comparison of H_1 (y-axis on the above plots) with QAOA $p = 1$ (x-axis on the above plots). The dashed purple line corresponds to equal performance.



(a) The approximation ratio for the QZ-inspired approach on 100 instances of SKM.

(b) The optimal time for the QZ-inspired approach on 100 instances of the SKM. The norm of each Hamiltonian, for each problem size has been fixed, according to Eq. 7.49, ensuring a fair comparison.

Figure B.4: Performance on the QZ-inspired approach on 100 instances of the SKM. The x-axis label refers to the order of T in the expansion of Eq. 7.55, with 0 being H_1 and e referring to the full exponential (i.e. Eq. 7.54).

Finally, in Fig. B.4 we assess the performance of the Quantum-Zermelo inspired approach for 100 SKM instances. Again, all the QZ-inspired approaches provide an improvement on the original H_1 Hamiltonian, indexed by 0 in the figures. Going to first order achieves a substantial improvement as with the Max-cut instances. The optimal times for the QZ-inspired approach can be found in Fig. B.4b for the SKM.

In this section, we have demonstrated that the approaches inspired by Hamiltonians for optimal state-transfer operate qualitatively similar on SKM as they do on Max-cut.

B.2 Instances for which QAOA $p = 1$ outperforms H_1

The aim of this section is to explore the instances where QAOA $p = 1$ has a shorter run-time than H_1 and/or provides a better approximation ratio, as shown in Fig. 7.11. The first thing to note is that these tend to be the exception, rather than the rule.

Looking first at the Max-cut instances for three-regular graphs, H_1 always provided a better approximation ratio than QAOA $p = 1$, typically in a much shorter time. There are four instances (two with the same approximation ratio) where QAOA has a similar or shorter run-time as highlighted in Fig. B.5. The corresponding graphs are shown in Table B.1. The first thing to note is that these problems are small, the largest

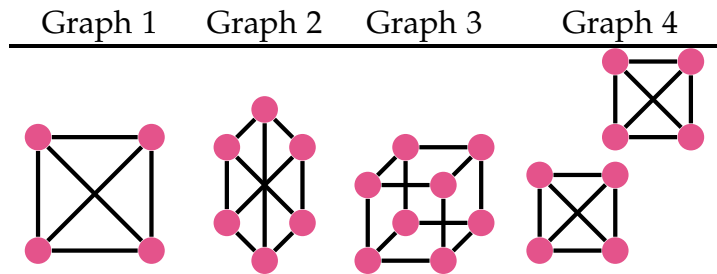


Table B.1: The exceptions for the three-regular graphs.

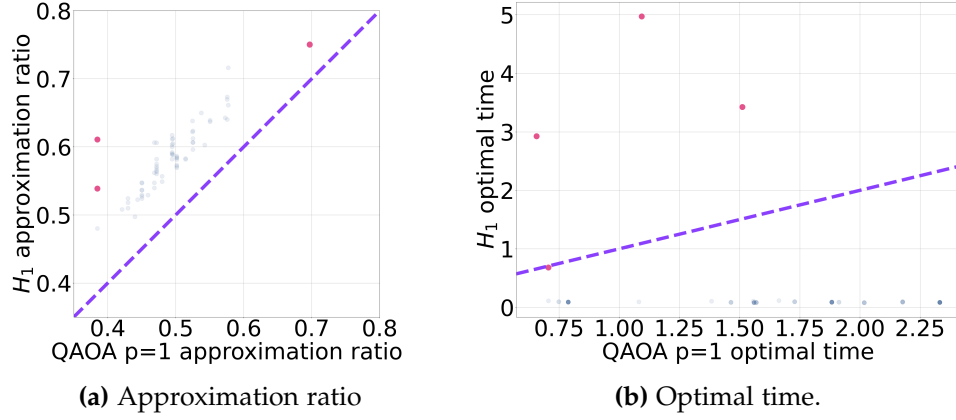


Figure B.5: The performance of H_1 compared to QAOA $p = 1$ on Max-cut with three-regular graphs. The instances with atypical run-times for H_1 are highlighted in pink. The rest of the data has been faded for clarity and to show how the remaining data is clustered, with darker regions corresponding to more data points. The dashed purple line shows both approaches performing the same.

being 8 qubits, despite problem sizes up to 12 qubits being considered. Graph 1 is a complete graph with four nodes, and Graph 4 is two copies of this graph. Both Graph 2 and 3 (a cube) consist of numerous small loops. Due to the relatively high degree of connectivity in these problems, it is likely H_1 is no longer operating in the local regime, as with the rest of the three-regular problems. In Fig. B.6 we investigate operating H_1 suboptimally, optimising only over run-times shorter than QAOA $p = 1$ for these four problems. The result is a negligible decrease in performance. Hence, even for these instances for which H_1 has atypical optimal times, it is possible for H_1 to provide a better approximation ratio in a shorter time than QAOA $p = 1$.

As with the three-regular graphs, H_1 always gave a better approximation ratio than QAOA $p = 1$ on Max-cut with the randomly generated graphs (highlighted in Fig. B.7). Similarly, we can look at the instances with atypical optimal times for H_1 . The story is similar to before, with only 116 out of 900 instances having run-times longer than QAOA. None of these problem instances consisted of more than 7 qubits (despite simulations going up to 12 qubits). Again, we conclude that these atypical optimal times are likely a small problem-setting phenomenon. We can

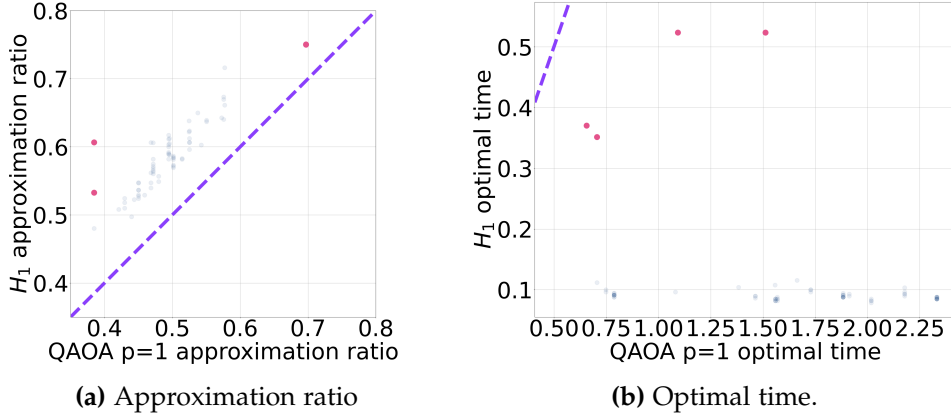


Figure B.6: The performance of H_1 compared to QAOA $p = 1$ on Max-cut with three-regular graphs. The highlighted pink dots show the instances with atypical H_1 run-times, optimised to give the best possible approximation ratio with shorter run-times than QAOA $p = 1$. The new run-times are shown in the lower plot, with the updated approximation ratio plotted in the upper plot. The dashed purple line shows both approaches performing the same.

also look for run-times of H_1 that provide a better approximation ratio in a shorter time than QAOA $p = 1$ for these problems. The results are shown in Fig. B.8 which shows the performance and new run-times of H_1 compared to QAOA $p = 1$. For all problem instances, it is possible to operate H_1 with a shorter run-time than QAOA $p = 1$ and provide a better approximation ratio. Unlike the previous discussion with three-regular graphs, the change in approximation ratio with these new run-times is not negligible for some of the problem instances.

Finally, we turn to the SKM instances, where QAOA $p = 1$ was able to provide a better approximation ratio than H_1 for 7 problem instances (out of 900). Five of these instances are 4-qubit problems, the remaining two are 5-qubit problems. In Fig. B.9 we have plotted the SKM data, including only those instances with problem sizes between 6 and 12 qubits. As we can see by ignoring small problem sizes from the data set, the behaviour is more predictable, with both the approximation ratio and the optimal time more clustered, largely independent of problem size.

In summary, we have demonstrated numerically that QAOA $p = 1$

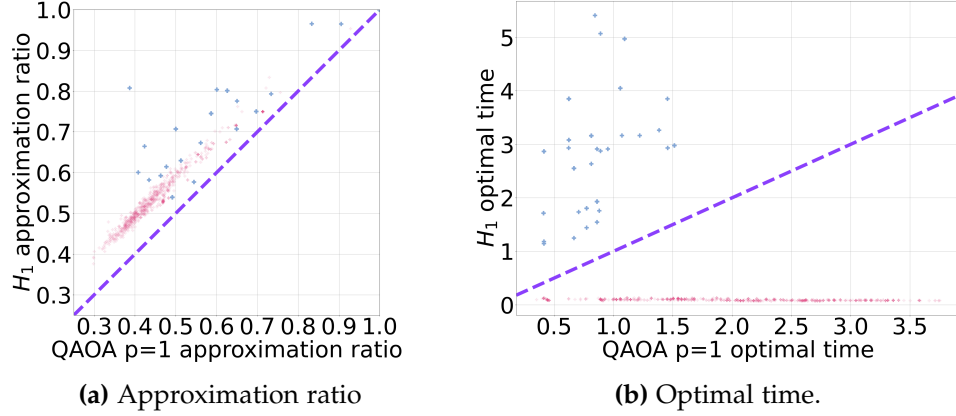


Figure B.7: The performance of H_1 compared to QAOA $p = 1$ on Max-cut with randomly generated graphs. The instances with atypical run-times for H_1 are highlighted in blue. The rest of the data has been faded for clarity and to show how the remaining data is clustered, with darker regions corresponding to more data points. The dashed purple line shows both approaches performing the same.

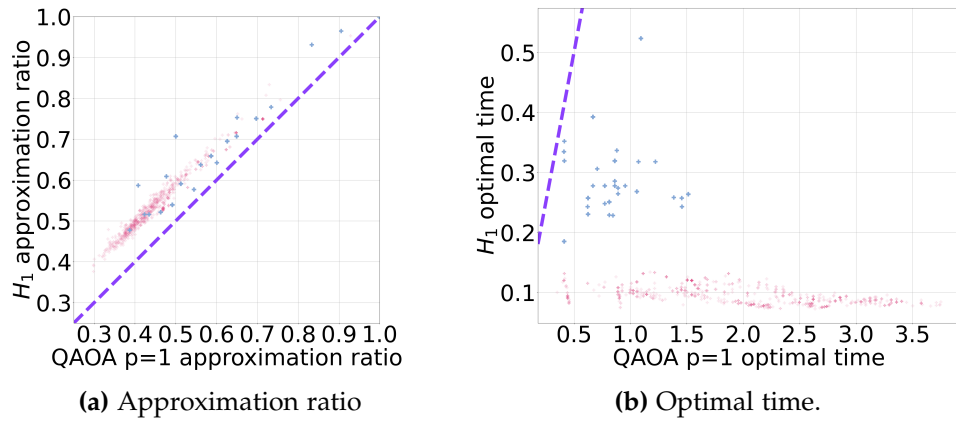


Figure B.8: The performance of H_1 compared to QAOA $p = 1$ on Max-cut with randomly generated graphs. The instances with atypical run-times for H_1 are highlighted in blue, these have been optimised to give the best possible approximation ratio with a corresponding run-time smaller than the QAOA $p = 1$ optimal time. The new run-times are shown in the lower plot, with the updated approximation ratio plotted in the upper plot. The rest of the data has been faded for clarity. The dashed purple line shows both approaches performing the same.

B. FURTHER NUMERICAL WORK ON RAPID HAMILTONIANS FOR OPTIMISATION

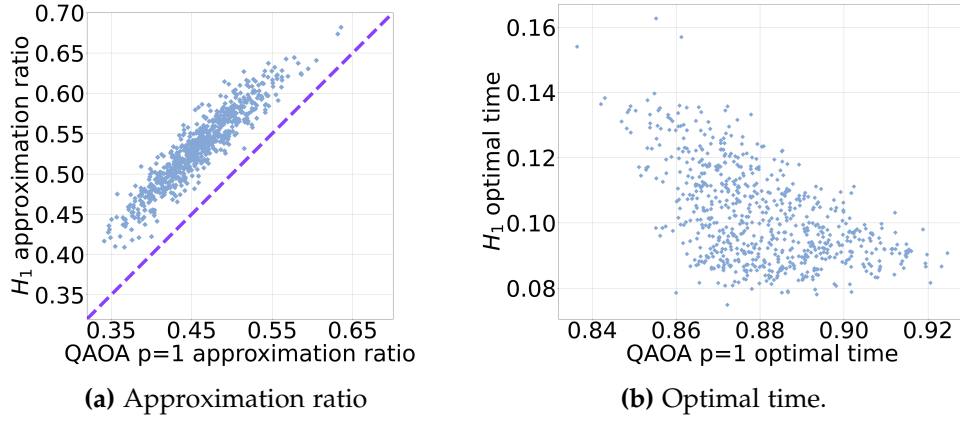


Figure B.9: The performance of H_1 compared to QAOA $p = 1$ on the SKM. Here we neglect smaller problem instances, plotting problem sizes between 6 and 12 qubits.

might have some advantages over H_1 on small, highly connected problems. In general, these problems are unlikely to be of any practical interest. Indeed, for the Max-cut instances, it was possible to operate H_1 suboptimally so that it still outperformed QAOA $p = 1$ with a shorter run-time.

Bibliography

- [1] L. D'ALESSIO, Y. KAFRI, A. POLKOVNIKOV, AND M. RIGOL, *From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics*, *Advances in Physics*, 65 (2016), pp. 239–362, [doi:10.1080/00018732.2016.1198134](https://doi.org/10.1080/00018732.2016.1198134).
- [2] T. KADOWAKI AND H. NISHIMORI, *Quantum annealing in the transverse Ising model*, *Phys. Rev. E*, 58 (1998), pp. 5355–5363, [doi:10.1103/PhysRevE.58.5355](https://doi.org/10.1103/PhysRevE.58.5355).
- [3] E. FARHI AND S. GUTMANN, *Quantum computation and decision trees*, *Phys. Rev. A*, 58 (1998), pp. 915–928, [doi:10.1103/PhysRevA.58.915](https://doi.org/10.1103/PhysRevA.58.915).
- [4] J. PRESKILL, *Quantum computing in the NISQ era and beyond*, *Quantum*, 2 (2018), p. 79, [doi:10.22331/q-2018-08-06-79](https://doi.org/10.22331/q-2018-08-06-79).
- [5] L. K. GROVER, *Quantum mechanics helps in searching for a needle in a haystack*, *Phys. Rev. Lett.*, 79 (1997), pp. 325–328, [doi:10.1103/PhysRevLett.79.325](https://doi.org/10.1103/PhysRevLett.79.325).
- [6] C. ZALKA, *Grover's quantum searching algorithm is optimal*, *Phys. Rev. A*, 60 (1999), pp. 2746–2751, [doi:10.1103/PhysRevA.60.2746](https://doi.org/10.1103/PhysRevA.60.2746).
- [7] C. DURR AND P. HOYER, *A quantum algorithm for finding the minimum*, *arXiv*, (1999), [doi:/10.48550/arXiv.quant-ph/9607014](https://doi.org/10.48550/arXiv.quant-ph/9607014).
- [8] A. GILLIAM, S. WOERNER, AND C. GONCIULEA, *Grover adaptive search for constrained polynomial binary optimization*, *Quantum*, 5 (2021), p. 428, [doi:10.22331/q-2021-04-08-428](https://doi.org/10.22331/q-2021-04-08-428).

- [9] E. M. STOUDENMIRE AND X. WAIN TAL, *Grover's algorithm offers no quantum advantage*, arXiv, (2023), [doi:10.48550/arXiv.2303.11317](https://doi.org/10.48550/arXiv.2303.11317).
- [10] E. FARHI, J. GOLDSTONE, S. GUTMANN, AND M. SIPSER, *Quantum computation by adiabatic evolution*, (2000), [doi:10.48550/ARXIV.QUANT-PH/0001106](https://doi.org/10.48550/ARXIV.QUANT-PH/0001106).
- [11] M. B. HASTINGS, *The power of adiabatic quantum computation with no sign problem*, Quantum, 5 (2021), p. 597, [doi:10.22331/q-2021-12-06-597](https://doi.org/10.22331/q-2021-12-06-597).
- [12] R. D. SOMMA, D. NAGAJ, AND M. KIEFEROVÁ, *Quantum speedup by quantum annealing*, Phys. Rev. Lett., 109 (2012), p. 050501, [doi:10.1103/PhysRevLett.109.050501](https://doi.org/10.1103/PhysRevLett.109.050501).
- [13] S. MUTHUKRISHNAN, T. ALBASH, AND D. A. LIDAR, *Sensitivity of quantum speedup by quantum annealing to a noisy oracle*, Phys. Rev. A, 99 (2019), p. 032324, [doi:10.1103/PhysRevA.99.032324](https://doi.org/10.1103/PhysRevA.99.032324).
- [14] E. FARHI, J. GOLDSTONE, AND S. GUTMANN, *A quantum approximate optimization algorithm*, arXiv, (2014), [doi:10.48550/arXiv.1411.4028](https://doi.org/10.48550/arXiv.1411.4028).
- [15] C. C. MCGEOCH AND P. FARRE, *Milestones on the quantum utility highway*, arXiv, (2023), [doi:10.48550/arXiv.2305.00883](https://doi.org/10.48550/arXiv.2305.00883).
- [16] A. D. KING, A. NOCERA, AND M. M. RAMS ET AL., *Computational supremacy in quantum simulation*, arXiv, (2024), [doi:10.48550/arXiv.2403.00910](https://doi.org/10.48550/arXiv.2403.00910).
- [17] S. EBADI, A. KEESLING, AND M. CAIN ET AL., *Quantum optimization of maximum independent set using Rydberg atom arrays*, Science, 376 (2022), pp. 1209–1215, [doi:10.1126/science.abo6587](https://doi.org/10.1126/science.abo6587).
- [18] B. ALTSHULER, H. KROVI, AND J. ROLAND, *Adiabatic quantum optimization fails for random instances of NP-complete problems*, arXiv, (2009), [doi:10.48550/arXiv.0908.2782](https://doi.org/10.48550/arXiv.0908.2782).

-
- [19] B. ALTSHULER, H. KROVI, AND J. ROLAND, *Anderson localization makes adiabatic quantum optimization fail*, Proceedings of the National Academy of Sciences, 107 (2010), pp. 12446–12450, [doi:10.1073/pnas.1002116107](https://doi.org/10.1073/pnas.1002116107).
- [20] D. WECKER, M. B. HASTINGS, AND M. TROYER, *Training a quantum optimizer*, Phys. Rev. A, 94 (2016), p. 022309, [doi:10.1103/PhysRevA.94.022309](https://doi.org/10.1103/PhysRevA.94.022309).
- [21] C. R. LAUMANN, R. MOESSNER, A. SCARDICCHIO, AND S. L. SONDHI, *Quantum adiabatic algorithm and scaling of gaps at first-order quantum phase transitions*, Phys. Rev. Lett., 109 (2012), p. 030502, [doi:10.1103/PhysRevLett.109.030502](https://doi.org/10.1103/PhysRevLett.109.030502).
- [22] A. CALLISON, N. CHANCELLOR, F. MINTERT, AND V. KENDON, *Finding spin glass ground states using quantum walks*, New Journal of Physics, 21 (2019), p. 123022, [doi:10.1088/1367-2630/ab5ca2](https://doi.org/10.1088/1367-2630/ab5ca2).
- [23] M. SREDNICKI, *Chaos and quantum thermalization*, Phys. Rev. E, 50 (1994), pp. 888–901, [doi:10.1103/PhysRevE.50.888](https://doi.org/10.1103/PhysRevE.50.888).
- [24] J. M. DEUTSCH, *Quantum statistical mechanics in a closed system*, Phys. Rev. A, 43 (1991), pp. 2046–2049, [doi:10.1103/PhysRevA.43.2046](https://doi.org/10.1103/PhysRevA.43.2046).
- [25] LABA AND TKACHUK, *Geometric characteristics of quantum evolution: curvature and torsion*, Condensed Matter Physics, 20 (2017), p. 13003, [doi:10.5488/cmp.20.13003](https://doi.org/10.5488/cmp.20.13003).
- [26] S. GOLDSTEIN, T. HARA, AND H. TASAKI, *The second law of thermodynamics for pure quantum states*, arXiv, (2013), [doi:10.48550/arXiv.1303.6393](https://doi.org/10.48550/arXiv.1303.6393).
- [27] C. ITOI AND M. AMANO, *The second law of thermodynamics from concavity of energy eigenvalues*, Journal of the Physical Society of Japan, 89 (2020), p. 104001, [doi:10.7566/jpsj.89.104001](https://doi.org/10.7566/jpsj.89.104001).

- [28] A. D. KING, S. SUZUKI, AND J. RAYMOND ET AL., *Coherent quantum annealing in a programmable 2,000 qubit Ising chain*, *Nature Physics*, 18 (2022), pp. 1324–1328, [doi:10.1038/s41567-022-01741-6](https://doi.org/10.1038/s41567-022-01741-6).
- [29] E. GIBNEY, *D-Wave upgrade: How scientists are using the world's most controversial quantum computer*, *Nature*, 541 (2017), pp. 447–448, [doi:10.1038/541447b](https://doi.org/10.1038/541447b).
- [30] C. H. PAPADIMITRIOU AND K. STEIGLITZ, *Combinatorial Optimization: Algorithms and Complexity*, Dover Publications, 1981.
- [31] W. SAKULER, J. M. OBERREUTER, R. AIOLFI, L. ASPRONI, B. ROMAN, AND J. SCHIEFER, *A real world test of portfolio optimization with quantum annealing*, *arXiv*, (2023), [doi:10.48550/arXiv.2303.12601](https://doi.org/10.48550/arXiv.2303.12601).
- [32] E. GRANT, T. S. HUMBLE, AND B. STUMP, *Benchmarking quantum annealing controls with portfolio optimization*, *Phys. Rev. Appl.*, 15 (2021), p. 014012, [doi:10.1103/PhysRevApplied.15.014012](https://doi.org/10.1103/PhysRevApplied.15.014012).
- [33] D. VENTURELLI AND A. KONDRATYEV, *Reverse quantum annealing approach to portfolio optimization problems*, *Quantum Machine Intelligence*, 1 (2019), p. 17–30, [doi:10.1007/s42484-019-00001-w](https://doi.org/10.1007/s42484-019-00001-w).
- [34] S. J. WEINBERG, F. SANCHES, T. IDE, K. KAMIYA, AND R. CORRELL, *Supply chain logistics with quantum and classical annealing algorithms*, *arXiv*, (2022), [doi:10.48550/arXiv.2205.04435](https://doi.org/10.48550/arXiv.2205.04435).
- [35] G. COLUCCI, S. V. D. LINDE, AND F. PHILLIPSON, *Power network optimization: A quantum approach*, *IEEE Access*, 11 (2023), pp. 98926–98938, [doi:10.1109/ACCESS.2023.3312997](https://doi.org/10.1109/ACCESS.2023.3312997).
- [36] R. K. NATH, H. THAPLIYAL, AND T. S. HUMBLE, *A review of machine learning classification using quantum annealing for real-world applications*, *SN Comput. Sci.*, 2 (2021), [doi:10.1007/s42979-021-00751-0](https://doi.org/10.1007/s42979-021-00751-0).
- [37] R. AYANZADEH, M. HALEM, AND T. FININ, *Reinforcement quantum annealing: A hybrid quantum learning automata*, *Scientific Reports*, 10 (2020), p. 7952, [doi:10.1038/s41598-020-64078-1](https://doi.org/10.1038/s41598-020-64078-1).

-
- [38] A. IRBÄCK, L. KNUTHSON, S. MOHANTY, AND C. PETERSON, *Folding lattice proteins with quantum annealing*, *Physical Review Research*, 4 (2022), [doi:10.1103/physrevresearch.4.043013](https://doi.org/10.1103/physrevresearch.4.043013).
- [39] K. MATO, R. MENGONI, D. OTTAVIANI, AND G. PALERMO, *Quantum molecular unfolding*, *Quantum Science and Technology*, 7 (2022), p. 035020, [doi:10.1088/2058-9565/ac73af](https://doi.org/10.1088/2058-9565/ac73af).
- [40] S. PURI, C. K. ANDERSEN, A. L. GRIMSMO, AND A. BLAIS, *Quantum annealing with all-to-all connected nonlinear oscillators*, *Nature Communications*, 8 (2017), p. 15785, [doi:10.1038/ncomms15785](https://doi.org/10.1038/ncomms15785).
- [41] A. LUCAS, *Ising formulations of many NP problems*, *Frontiers in Physics*, 2 (2014), [doi:10.3389/fphy.2014.00005](https://doi.org/10.3389/fphy.2014.00005).
- [42] I. HEN AND F. M. SPEDALIERI, *Quantum annealing for constrained optimization*, *Physical Review Applied*, 5 (2016), [doi:10.1103/physrevapplied.5.034007](https://doi.org/10.1103/physrevapplied.5.034007).
- [43] I. HEN AND M. S. SARANDY, *Driver hamiltonians for constrained optimization in quantum annealing*, *Physical Review A*, 93 (2016), [doi:10.1103/physreva.93.062312](https://doi.org/10.1103/physreva.93.062312).
- [44] N. CHANCELLOR, *Domain wall encoding of discrete variables for quantum annealing and QAOA*, *Quantum Science and Technology*, 4 (2019), p. 045004, [doi:10.1088/2058-9565/ab33c2](https://doi.org/10.1088/2058-9565/ab33c2).
- [45] K. M. ZICK, O. SHEHAB, AND M. FRENCH, *Experimental quantum annealing: case study involving the graph isomorphism problem*, *Scientific Reports*, 5 (2015), [doi:10.1038/srep11168](https://doi.org/10.1038/srep11168).
- [46] A. BRAIDA, S. MARTIEL, AND I. TODINCA, *On constant-time quantum annealing and guaranteed approximations for graph optimization problems*, *Quantum Science and Technology*, 7 (2022), p. 045030, [doi:10.1088/2058-9565/ac8e91](https://doi.org/10.1088/2058-9565/ac8e91).

- [47] D. SHERRINGTON AND S. KIRKPATRICK, *Solvable model of a spin-glass*, Phys. Rev. Lett., 35 (1975), pp. 1792–1796, [doi:10.1103/PhysRevLett.35.1792](https://doi.org/10.1103/PhysRevLett.35.1792).
- [48] T. ALBASH AND D. A. LIDAR, *Adiabatic quantum computation*, Reviews of Modern Physics, 90 (2018), [doi:10.1103/revmodphys.90.015002](https://doi.org/10.1103/revmodphys.90.015002).
- [49] A. ABBAS, A. AMBAINIS, AND B. AUGUSTINO ET AL., *Quantum optimization: Potential, challenges, and the path forward*, 2023. <https://arxiv.org/abs/2312.02279>.
- [50] M. GAREY, D. JOHNSON, AND L. STOCKMEYER, *Some simplified NP-complete graph problems*, Theoretical Computer Science, 1 (1976), pp. 237–267, [doi:https://doi.org/10.1016/0304-3975\(76\)90059-1](https://doi.org/10.1016/0304-3975(76)90059-1).
- [51] C. H. PAPADIMITRIOU AND M. YANNAKAKIS, *Optimization, approximation, and complexity classes*, Journal of Computer and System Sciences, 43 (1991), pp. 425–440, [doi:https://doi.org/10.1016/0022-0000\(91\)90023-X](https://doi.org/10.1016/0022-0000(91)90023-X).
- [52] V. S. DENCHEV, S. BOIXO, S. V. ISAKOV, N. DING, R. BABBUSH, V. SMELYANSKIY, J. MARTINIS, AND H. NEVEN, *What is the computational value of finite-range tunneling?*, Phys. Rev. X, 6 (2016), p. 031015, [doi:10.1103/PhysRevX.6.031015](https://doi.org/10.1103/PhysRevX.6.031015).
- [53] H. M. BAUZA AND D. A. LIDAR, *Scaling advantage in approximate optimization with quantum annealing*, arXiv, (2024), [doi:10.48550/arXiv.2401.07184](https://doi.org/10.48550/arXiv.2401.07184).
- [54] T. KADOWAKI AND H. NISHIMORI, *Greedy parameter optimization for diabatic quantum annealing*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 381 (2022), [doi:10.1098/rsta.2021.0416](https://doi.org/10.1098/rsta.2021.0416).

-
- [55] J. ROLAND AND N. J. CERF, *Quantum search by local adiabatic evolution*, Phys. Rev. A, 65 (2002), p. 042308, [doi:10.1103/PhysRevA.65.042308](https://doi.org/10.1103/PhysRevA.65.042308).
- [56] A. DEL CAMPO, *Shortcuts to adiabaticity by counterdiabatic driving*, Phys. Rev. Lett., 111 (2013), p. 100502, [doi:10.1103/PhysRevLett.111.100502](https://doi.org/10.1103/PhysRevLett.111.100502).
- [57] D. SELS AND A. POLKOVNIKOV, *Minimizing irreversible losses in quantum systems by local counterdiabatic driving*, Proceedings of the National Academy of Sciences, 114 (2017), pp. E3909–E3916, [doi:10.1073/pnas.1619826114](https://doi.org/10.1073/pnas.1619826114).
- [58] F. PETIZIOL, B. DIVE, F. MINTERT, AND S. WIMBERGER, *Fast adiabatic evolution by oscillating initial hamiltonians*, Phys. Rev. A, 98 (2018), p. 043436, [doi:10.1103/PhysRevA.98.043436](https://doi.org/10.1103/PhysRevA.98.043436).
- [59] L. PRIELINGER, A. HARTMANN, Y. YAMASHIRO, K. NISHIMURA, W. LECHNER, AND H. NISHIMORI, *Two-parameter counter-diabatic driving in quantum annealing*, Phys. Rev. Res., 3 (2021), p. 013227, [doi:10.1103/PhysRevResearch.3.013227](https://doi.org/10.1103/PhysRevResearch.3.013227).
- [60] I. ČEPAITĖ, A. POLKOVNIKOV, A. J. DALEY, AND C. W. DUNCAN, *Counterdiabatic optimized local driving*, PRX Quantum, 4 (2023), p. 010312, [doi:10.1103/PRXQuantum.4.010312](https://doi.org/10.1103/PRXQuantum.4.010312).
- [61] M. H. S. AMIN, P. J. LOVE, AND C. J. S. TRUNCIK, *Thermally assisted adiabatic quantum computation*, Physical Review Letters, 100 (2008), [doi:10.1103/physrevlett.100.060503](https://doi.org/10.1103/physrevlett.100.060503).
- [62] N. G. DICKSON, M. W. JOHNSON, AND M. H. AMIN ET AL., *Thermally assisted quantum annealing of a 16-qubit problem*, Nature Communications, 4 (2013), p. 1903, [doi:10.1038/ncomms2920](https://doi.org/10.1038/ncomms2920).
- [63] E. J. CROSSON AND D. A. LIDAR, *Prospects for quantum enhancement with diabatic quantum annealing*, Nature Reviews Physics, 3 (2021), pp. 466–489, [doi:10.1038/s42254-021-00313-6](https://doi.org/10.1038/s42254-021-00313-6).

- [64] H. G. KATZGRABER, F. HAMZE, Z. ZHU, A. J. OCHOA, AND H. MUNOZ-BAUZA, *Seeking quantum speedup through spin glasses: The good, the bad, and the ugly*, Physical Review X, 5 (2015), [doi:10.1103/physrevx.5.031026](https://doi.org/10.1103/physrevx.5.031026).
- [65] A. BRAIDA, S. MARTIEL, AND I. TODINCA, *Tight Lieb–Robinson bound for approximation ratio in quantum annealing*, npj Quantum Information, 10 (2024), p. 40, [doi:10.1038/s41534-024-00832-x](https://doi.org/10.1038/s41534-024-00832-x).
- [66] A. G. DE OLIVEIRA, E. DIAMOND-HITCHCOCK, D. M. WALKER, M. T. WELLS-PESTELL, G. PELEGRÍ, C. J. PICKEN, G. P. A. MALCOLM, A. J. DALEY, J. BASS, AND J. D. PRITCHARD, *Demonstration of weighted graph optimization on a Rydberg atom array using local light-shifts*, arXiv, (2024), [doi:10.48550/arXiv.2404.02658](https://doi.org/10.48550/arXiv.2404.02658).
- [67] A. CALLISON, M. FESTENSTEIN, J. CHEN, L. NITA, V. KENDON, AND N. CHANCELLOR, *Energetic perspective on rapid quenches in quantum annealing*, PRX Quantum, 2 (2021), p. 010338, [doi:10.1103/PRXQuantum.2.010338](https://doi.org/10.1103/PRXQuantum.2.010338).
- [68] A. CALLISON, *Continuous-time Quantum Computing*, PhD thesis, Imperial College London, 2021, [doi:10.25560/91503](https://doi.org/10.25560/91503).
- [69] A. M. CHILDS AND J. GOLDSTONE, *Spatial search by quantum walk*, Phys. Rev. A, 70 (2004), p. 022314, [doi:10.1103/PhysRevA.70.022314](https://doi.org/10.1103/PhysRevA.70.022314).
- [70] N. CHANCELLOR, *Modernizing quantum annealing using local searches*, New Journal of Physics, 19 (2017), p. 023024, [doi:10.1088/1367-2630/aa59c4](https://doi.org/10.1088/1367-2630/aa59c4).
- [71] N. CHANCELLOR, *Modernizing quantum annealing ii: genetic algorithms with the inference primitive formalism*, Natural Computing, 22 (2023), pp. 737–752, [doi:10.1007/s11047-022-09905-2](https://doi.org/10.1007/s11047-022-09905-2).

-
- [72] N. CHANCELLOR AND V. KENDON, *Experimental test of search range in quantum annealing*, Phys. Rev. A, 104 (2021), p. 012604, [doi:10.1103/PhysRevA.104.012604](https://doi.org/10.1103/PhysRevA.104.012604).
- [73] H. WANG, H.-C. YEH, AND A. KAMENEV, *Many-body localization enables iterative quantum optimization*, Nature Communications, 13 (2022), p. 5503, [doi:10.1038/s41467-022-33179-y](https://doi.org/10.1038/s41467-022-33179-y).
- [74] H. ZHANG, K. BOOTHBY, AND A. KAMENEV, *Cyclic quantum annealing: Searching for deep low-energy states in 5000-qubit spin glass*, arXiv, (2024), [doi:10.48550/arXiv.2403.01034](https://doi.org/10.48550/arXiv.2403.01034).
- [75] A. PERUZZO, J. MCCLEAN, P. SHADBOLT, M.-H. YUNG, X.-Q. ZHOU, P. J. LOVE, A. ASPURU-GUZI, AND J. L. O'BRIEN, *A variational eigenvalue solver on a photonic quantum processor*, Nature Communications, 5 (2014), p. 4213, [doi:10.1038/ncomms5213](https://doi.org/10.1038/ncomms5213).
- [76] L. ZHOU, S.-T. WANG, S. CHOI, H. PICHLER, AND M. D. LUKIN, *Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices*, Phys. Rev. X, 10 (2020), p. 021067, [doi:10.1103/PhysRevX.10.021067](https://doi.org/10.1103/PhysRevX.10.021067).
- [77] M. P. HARRIGAN, K. J. SUNG, M. NEELEY, AND K. J. SATZINGER ET AL., *Quantum approximate optimization of non-planar graph problems on a planar superconducting processor*, Nature Physics, 17 (2021), pp. 332–336, [doi:10.1038/s41567-020-01105-y](https://doi.org/10.1038/s41567-020-01105-y).
- [78] T. M. GRAHAM, Y. SONG, AND J. SCOTT ET AL., *Multi-qubit entanglement and algorithms on a neutral-atom quantum computer*, Nature, 604 (2022), pp. 457–462, [doi:10.1038/s41586-022-04603-6](https://doi.org/10.1038/s41586-022-04603-6).
- [79] J. S. OTTERBACH, R. MANENTI, AND N. ALIDOUST ET AL., *Unsupervised machine learning on a hybrid quantum computer*, arXiv, (2017), [doi:10.48550/ARXIV.1712.05771](https://doi.org/10.48550/ARXIV.1712.05771).

- [80] Z. WANG, S. HADFIELD, Z. JIANG, AND E. G. RIEFFEL, *Quantum approximate optimization algorithm for MaxCut: A fermionic view*, Physical Review A, 97 (2018), [doi:10.1103/physreva.97.022304](https://doi.org/10.1103/physreva.97.022304).
- [81] G. B. MBENG, R. FAZIO, AND G. SANTORO, *Quantum annealing: a journey through digitalization, control, and hybrid quantum variational schemes*, arXiv, (2019), [doi:10.48550/ARXIV.1906.08948](https://doi.org/10.48550/ARXIV.1906.08948).
- [82] M. A. NIELSEN AND I. L. CHUANG, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, Cambridge University Press, 2010.
- [83] R. ACHARYA, I. ALEINER, AND R. ALLEN ET AL., *Suppressing quantum errors by scaling a surface code logical qubit*, Nature, 614 (2023), pp. 676–681, [doi:10.1038/s41586-022-05434-1](https://doi.org/10.1038/s41586-022-05434-1).
- [84] B. F. SCHIFFER, A. F. RUBIO, R. TRIVEDI, AND J. I. CIRAC, *The quantum adiabatic algorithm suppresses the proliferation of errors*, arXiv, (2024), [doi:10.48550/arXiv.2404.15397](https://doi.org/10.48550/arXiv.2404.15397).
- [85] R. TRIVEDI, A. FRANCO RUBIO, AND J. I. CIRAC, *Quantum advantage and stability to errors in analogue quantum simulators*, Nature Communications, 15 (2024), p. 6507, [doi:10.1038/s41467-024-50750-x](https://doi.org/10.1038/s41467-024-50750-x).
- [86] A. M. CHILDS, E. FARHI, AND J. PRESKILL, *Robustness of adiabatic quantum computation*, Physical Review A, 65 (2001), [doi:10.1103/physreva.65.012322](https://doi.org/10.1103/physreva.65.012322).
- [87] J. GONZALEZ-CONDE, Z. MORRELL, M. VUFFRAY, T. ALBASH, AND C. COFFRIN, *The cost of emulating a small quantum annealing problem in the circuit-model*, arXiv, (2024), [doi:10.48550/arXiv.2402.17667](https://doi.org/10.48550/arXiv.2402.17667).
- [88] E. PELOFSKE, A. BÄRTSCHI, AND S. EIDENBENZ, *Short-depth QAOA circuits and quantum annealing on higher-order Ising models*, npj Quantum Information, 10 (2024), p. 30, [doi:10.1038/s41534-024-00825-w](https://doi.org/10.1038/s41534-024-00825-w).

-
- [89] S. HADFIELD, Z. WANG, B. O'GORMAN, E. RIEFFEL, D. VENTURELLI, AND R. BISWAS, *From the quantum approximate optimization algorithm to a quantum alternating operator ansatz*, *Algorithms*, 12 (2019), p. 34, [doi:10.3390/a12020034](https://doi.org/10.3390/a12020034).
- [90] D. J. EGGER, J. MAREČEK, AND S. WOERNER, *Warm-starting quantum optimization*, *Quantum*, 5 (2021), p. 479, [doi:10.22331/q-2021-06-17-479](https://doi.org/10.22331/q-2021-06-17-479).
- [91] J. R. MCCLEAN, S. BOIXO, V. N. SMELYANSKIY, R. BABBUSH, AND H. NEVEN, *Barren plateaus in quantum neural network training landscapes*, *Nature Communications*, 9 (2018), p. 4812, [doi:10.1038/s41467-018-07090-4](https://doi.org/10.1038/s41467-018-07090-4).
- [92] L. BITTEL AND M. KLIESCH, *Training variational quantum algorithms is NP-Hard*, *Physical Review Letters*, 127 (2021), [doi:10.1103/physrevlett.127.120502](https://doi.org/10.1103/physrevlett.127.120502).
- [93] D. AMARO, C. MODICA, M. ROSENKRANZ, M. FIORENTINI, M. BENEDETTI, AND M. LUBASCH, *Filtering variational quantum algorithms for combinatorial optimization*, *Quantum Science and Technology*, 7 (2022), p. 015021, [doi:10.1088/2058-9565/ac3e54](https://doi.org/10.1088/2058-9565/ac3e54).
- [94] K. BLEKOS, D. BRAND, A. Ceschini, C.-H. CHOU, R.-H. LI, K. PANDYA, AND A. SUMMER, *A review on quantum approximate optimization algorithm and its variants*, *Physics Reports*, 1068 (2024), p. 1–66, [doi:10.1016/j.physrep.2024.03.002](https://doi.org/10.1016/j.physrep.2024.03.002).
- [95] L. T. BRADY, C. L. BALDWIN, A. BAPAT, Y. KHARKOV, AND A. V. GORSHKOV, *Optimal protocols in quantum annealing and quantum approximate optimization algorithm problems*, *Phys. Rev. Lett.*, 126 (2021), p. 070505, [doi:10.1103/PhysRevLett.126.070505](https://doi.org/10.1103/PhysRevLett.126.070505).
- [96] L. T. BRADY, L. KOCIA, P. BIENIAS, A. BAPAT, Y. KHARKOV, AND A. V. GORSHKOV, *Behavior of analog quantum algorithms*, *arXiv*, (2021), [doi:10.48550/ARXIV.2107.01218](https://doi.org/10.48550/ARXIV.2107.01218).

- [97] X. FEI, L. T. BRADY, J. LARSON, S. LEYFFER, AND S. SHEN, *Binary control pulse optimization for quantum systems*, Quantum, 7 (2023), p. 892, [doi:10.22331/q-2023-01-04-892](https://doi.org/10.22331/q-2023-01-04-892).
- [98] L. C. VENUTI, D. D'ALESSANDRO, AND D. A. LIDAR, *Optimal control for quantum optimization of closed and open systems*, Physical Review Applied, 16 (2021), [doi:10.1103/physrevapplied.16.054023](https://doi.org/10.1103/physrevapplied.16.054023).
- [99] L. GERBLICH, T. DASANJH, H. Q. X. WONG, D. ROSS, L. NOVO, N. CHANCELLOR, AND V. KENDON, *Advantages of multistage quantum walks over QAOA*, arXiv, (2024), [doi:10.48550/arXiv.2407.06663](https://doi.org/10.48550/arXiv.2407.06663).
- [100] C. LIN, Y. WANG, G. KOLESOV, AND U. C. V. KALABIĆ, *Application of Pontryagin's minimum principle to Grover's quantum search problem*, Phys. Rev. A, 100 (2019), p. 022327, [doi:10.1103/PhysRevA.100.022327](https://doi.org/10.1103/PhysRevA.100.022327).
- [101] S. SCHULZ, D. WILLSCH, AND K. MICHELSSEN, *Guided quantum walk*, Physical Review Research, 6 (2024), [doi:10.1103/physrevresearch.6.013312](https://doi.org/10.1103/physrevresearch.6.013312).
- [102] J. D. HUNTER, *Matplotlib: A 2D graphics environment*, Computing in Science & Engineering, 9 (2007), pp. 90–95, [doi:10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [103] F. M. DEKKING, C. KRAAIKAMP, H. P. LOPUHAÄ, AND L. E. MEESTER, *A Modern Introduction to Probability and Statistics*, Springer London, 2005.
- [104] J. JOHANSSON, P. NATION, AND F. NORI, *QuTiP: An open-source Python framework for the dynamics of open quantum systems*, Computer Physics Communications, 183 (2012), pp. 1760–1772, [doi:10.1016/j.cpc.2012.02.021](https://doi.org/10.1016/j.cpc.2012.02.021).
- [105] J. JOHANSSON, P. NATION, AND F. NORI, *QuTiP 2: A Python framework for the dynamics of open quantum systems*, Com-

- puter Physics Communications, 184 (2013), pp. 1234–1240, [doi:https://doi.org/10.1016/j.cpc.2012.11.019](https://doi.org/10.1016/j.cpc.2012.11.019).
- [106] A. A. HAGBERG, D. A. SCHULT, AND P. J. SWART, *Exploring network structure, dynamics, and function using NetworkX*, in Proceedings of the 7th Python in Science Conference, G. Varoquaux, T. Vaught, and J. Millman, eds., Pasadena, CA USA, 2008, pp. 11 – 15.
- [107] H. CHEN AND D. A. LIDAR, *Hamiltonian open quantum system toolkit*, Communications Physics, 5 (2022), p. 112, [doi:10.1038/s42005-022-00887-2](https://doi.org/10.1038/s42005-022-00887-2).
- [108] S. J. BLUNDELL AND K. M. BLUNDELL, *Concepts in Thermal Physics*, Oxford University Press, 10 2009.
- [109] H. WILMING, T. R. DE OLIVEIRA, A. J. SHORT, AND J. EISERT, *Equilibration Times in Closed Quantum Many-Body Systems*, Springer International Publishing, Cham, 2018, pp. 435–455.
- [110] T. R. DE OLIVEIRA, C. CHARALAMBOUS, D. JONATHAN, M. LEWENSTEIN, AND A. RIERA, *Equilibration time scales in closed many-body quantum systems*, New Journal of Physics, 20 (2018), p. 033032, [doi:10.1088/1367-2630/aab03b](https://doi.org/10.1088/1367-2630/aab03b).
- [111] H. WILMING, M. GOIHL, C. KRUMNOW, AND J. EISERT, *Towards local equilibration in closed interacting quantum many-body systems*, arXiv, (2017), [doi:10.48550/arXiv.1704.06291](https://doi.org/10.48550/arXiv.1704.06291).
- [112] J. M. DEUTSCH, *Eigenstate thermalization hypothesis*, Reports on Progress in Physics, 81 (2018), p. 082001, [doi:10.1088/1361-6633/aac9f1](https://doi.org/10.1088/1361-6633/aac9f1).
- [113] J. R. GARRISON AND T. GROVER, *Does a single eigenstate encode the full hamiltonian?*, Physical Review X, 8 (2018), [doi:10.1103/physrevx.8.021026](https://doi.org/10.1103/physrevx.8.021026).

- [114] M. RIGOL, V. DUNJKO, AND M. OLSHANIL, *Thermalization and its mechanism for generic isolated quantum systems*, *Nature*, 452 (2008), pp. 854–858, [doi:10.1038/nature06838](https://doi.org/10.1038/nature06838).
- [115] F. H. L. ESSLER AND M. FAGOTTI, *Quench dynamics and relaxation in isolated integrable quantum spin chains*, *Journal of Statistical Mechanics: Theory and Experiment*, 2016 (2016), p. 064002, [doi:10.1088/1742-5468/2016/06/064002](https://doi.org/10.1088/1742-5468/2016/06/064002).
- [116] R. NANDKISHORE AND D. A. HUSE, *Many-body localization and thermalization in quantum statistical mechanics*, *Annual Review of Condensed Matter Physics*, 6 (2015), pp. 15–38, [doi:10.1146/annurev-conmatphys-031214-014726](https://doi.org/10.1146/annurev-conmatphys-031214-014726).
- [117] P. REIMANN, *Eigenstate thermalization: Deutsch's approach and beyond*, *New Journal of Physics*, 17 (2015), p. 055025, [doi:10.1088/1367-2630/17/5/055025](https://doi.org/10.1088/1367-2630/17/5/055025).
- [118] H. KIM, T. N. IKEDA, AND D. A. HUSE, *Testing whether all eigenstates obey the eigenstate thermalization hypothesis*, *Phys. Rev. E*, 90 (2014), p. 052105, [doi:10.1103/PhysRevE.90.052105](https://doi.org/10.1103/PhysRevE.90.052105).
- [119] J. M. DEUTSCH, *Eigenstate thermalization hypothesis*, *Reports on Progress in Physics*, 81 (2018), p. 082001, [doi:10.1088/1361-6633/aac9f1](https://doi.org/10.1088/1361-6633/aac9f1).
- [120] K. R. FRATUS AND M. SREDNICKI, *Eigenstate thermalization in systems with spontaneously broken symmetry*, *Phys. Rev. E*, 92 (2015), p. 040103, [doi:10.1103/PhysRevE.92.040103](https://doi.org/10.1103/PhysRevE.92.040103).
- [121] K. R. FRATUS AND M. A. SREDNICKI, *Eigenstate thermalization and spontaneous symmetry breaking in the one-dimensional transverse-field Ising model with power-law interactions*, *arXiv: Statistical Mechanics*, (2016), [doi:10.48550/arXiv.1611.03992](https://doi.org/10.48550/arXiv.1611.03992).

-
- [122] A. POLKOVNIKOV, *Microscopic diagonal entropy and its connection to basic thermodynamic relations*, *Annals of Physics*, 326 (2011), pp. 486–499, [doi:10.1016/j.aop.2010.08.004](https://doi.org/10.1016/j.aop.2010.08.004).
- [123] H. TASAKI, *The second law of thermodynamics as a theorem in quantum mechanics*, *arXiv*, (2000), [doi:10.48550/arXiv.cond-mat/0011321](https://doi.org/10.48550/arXiv.cond-mat/0011321).
- [124] T. N. IKEDA, N. SAKUMICHI, A. POLKOVNIKOV, AND M. UEDA, *The second law of thermodynamics under unitary evolution and external operations*, *Annals of Physics*, 354 (2015), p. 338–352, [doi:10.1016/j.aop.2015.01.003](https://doi.org/10.1016/j.aop.2015.01.003).
- [125] K. KANEKO, E. IYODA, AND T. SAGAWA, *Work extraction from a single energy eigenstate*, *Phys. Rev. E*, 99 (2019), p. 032128, [doi:10.1103/PhysRevE.99.032128](https://doi.org/10.1103/PhysRevE.99.032128).
- [126] N. KOUKOULEKIDIS, R. ALEXANDER, T. HEBDIGE, AND D. JENNINGS, *The geometry of passivity for quantum systems and a novel elementary derivation of the Gibbs state*, *Quantum*, 5 (2021), p. 411, [doi:10.22331/q-2021-03-15-411](https://doi.org/10.22331/q-2021-03-15-411).
- [127] P. SKRZYPCZYK, R. SILVA, AND N. BRUNNER, *Passivity, complete passivity, and virtual temperatures*, *Phys. Rev. E*, 91 (2015), p. 052133, [doi:10.1103/PhysRevE.91.052133](https://doi.org/10.1103/PhysRevE.91.052133).
- [128] W. W. HO, T. MORI, D. A. ABANIN, AND E. G. DALLA TORRE, *Quantum and classical Floquet prethermalization*, *Annals of Physics*, 454 (2023), p. 169297, [doi:10.1016/j.aop.2023.169297](https://doi.org/10.1016/j.aop.2023.169297).
- [129] D. HAHN, D. J. LUITZ, AND J. T. CHALKER, *Eigenstate correlations, the eigenstate thermalization hypothesis, and quantum information dynamics in chaotic many-body quantum systems*, *Phys. Rev. X*, 14 (2024), p. 031029, [doi:10.1103/PhysRevX.14.031029](https://doi.org/10.1103/PhysRevX.14.031029).
- [130] A. CHAN, A. DE LUCA, AND J. T. CHALKER, *Eigenstate correlations, thermalization, and the butterfly effect*, *Phys. Rev. Lett.*, 122 (2019), p. 220601, [doi:10.1103/PhysRevLett.122.220601](https://doi.org/10.1103/PhysRevLett.122.220601).

- [131] M. BRENES, S. PAPPALARDI, M. T. MITCHISON, J. GOOLD, AND A. SILVA, *Out-of-time-order correlations and the fine structure of eigenstate thermalization*, Phys. Rev. E, 104 (2021), p. 034120, [doi:10.1103/PhysRevE.104.034120](https://doi.org/10.1103/PhysRevE.104.034120).
- [132] D. L. JAFFERIS, D. K. KOLCHMEYER, B. MUKHAMETZHANOV, AND J. SONNER, *Matrix models for eigenstate thermalization*, Phys. Rev. X, 13 (2023), p. 031033, [doi:10.1103/PhysRevX.13.031033](https://doi.org/10.1103/PhysRevX.13.031033).
- [133] L. FOINI AND J. KURCHAN, *Eigenstate thermalization hypothesis and out of time order correlators*, Phys. Rev. E, 99 (2019), p. 042139, [doi:10.1103/PhysRevE.99.042139](https://doi.org/10.1103/PhysRevE.99.042139).
- [134] P. MIRKARIMI, A. CALLISON, L. LIGHT, N. CHANCELLOR, AND V. KENDON, *Comparing the hardness of MAX 2-SAT problem instances for quantum and classical algorithms*, Phys. Rev. Res., 5 (2023), p. 023151, [doi:10.1103/PhysRevResearch.5.023151](https://doi.org/10.1103/PhysRevResearch.5.023151).
- [135] J. WURTZ AND D. LYKOV, *Fixed-angle conjectures for the quantum approximate optimization algorithm on regular maxcut graphs*, Phys. Rev. A, 104 (2021), p. 052419, [doi:10.1103/PhysRevA.104.052419](https://doi.org/10.1103/PhysRevA.104.052419).
- [136] K. GNATENKO, H. LABA, AND V. TKACHUK, *Geometric properties of evolutionary graph states and their detection on a quantum computer*, Physics Letters A, 452 (2022), p. 128434, [doi:https://doi.org/10.1016/j.physleta.2022.128434](https://doi.org/10.1016/j.physleta.2022.128434).
- [137] T. PROSEN, *Ergodic properties of a generic nonintegrable quantum many-body system in the thermodynamic limit*, Phys. Rev. E, 60 (1999), pp. 3949–3968, [doi:10.1103/PhysRevE.60.3949](https://doi.org/10.1103/PhysRevE.60.3949).
- [138] R. MONDAINI AND M. RIGOL, *Eigenstate thermalization in the two-dimensional transverse field Ising model. ii. off-diagonal matrix elements of observables*, Phys. Rev. E, 96 (2017), p. 012157, [doi:10.1103/PhysRevE.96.012157](https://doi.org/10.1103/PhysRevE.96.012157).

-
- [139] W. BEUGELING, R. MOESSNER, AND M. HAQUE, *Off-diagonal matrix elements of local operators in many-body quantum systems*, Phys. Rev. E, 91 (2015), p. 012144, [doi:10.1103/PhysRevE.91.012144](https://doi.org/10.1103/PhysRevE.91.012144).
- [140] E. FARHI, D. GOSSET, I. HEN, A. W. SANDVIK, P. SHOR, A. P. YOUNG, AND F. ZAMPONI, *Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs*, Physical Review A, 86 (2012), [doi:10.1103/physreva.86.052334](https://doi.org/10.1103/physreva.86.052334).
- [141] M. JEANSONNE AND J. FOLEY, *Review of the Exponentially Modified Gaussian (emg) function since 1983*, Journal of Chromatographic Science, 29 (1991), pp. 258–266, [doi:10.1093/chromsci/29.6.258](https://doi.org/10.1093/chromsci/29.6.258).
- [142] Y. KALAMBET, Y. KOZMIN, K. MIKHAILOVA, I. NAGAEV, AND P. TIKHONOV, *Reconstruction of chromatographic peaks using the exponentially modified Gaussian function*, Journal of Chemometrics, 25 (2011), pp. 352–356, [doi:10.1002/cem.1343](https://doi.org/10.1002/cem.1343).
- [143] A. FELINGER, T. PAP, AND J. INCZÉDY, *Curve fitting to asymmetrical chromatograms by the extended Kalman filter in frequency domain*, Talanta, 41 (1994), pp. 1119–1126, [doi:10.1016/0039-9140\(94\)80081-2](https://doi.org/10.1016/0039-9140(94)80081-2).
- [144] K. F. RILEY, M. P. HOBSON, AND S. J. BENCE, *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*, Cambridge University Press, 3 ed., 2006.
- [145] E. CROSSON AND S. SLEZAK, *Classical simulation of high temperature quantum Ising models*, arXiv, (2020), [doi:10.48550/ARXIV.2002.02232](https://doi.org/10.48550/ARXIV.2002.02232).
- [146] J. M. DEUTSCH, *Thermodynamic entropy of a many-body energy eigenstate*, New Journal of Physics, 12 (2010), p. 075021, [doi:10.1088/1367-2630/12/7/075021](https://doi.org/10.1088/1367-2630/12/7/075021).
- [147] J. M. DEUTSCH, H. LI, AND A. SHARMA, *Microscopic origin of thermodynamic entropy in isolated systems*, Phys. Rev. E, 87 (2013), p. 042135, [doi:10.1103/PhysRevE.87.042135](https://doi.org/10.1103/PhysRevE.87.042135).

- [148] L. F. SANTOS, A. POLKOVNIKOV, AND M. RIGOL, *Entropy of isolated quantum systems after a quench*, Phys. Rev. Lett., 107 (2011), p. 040601, [doi:10.1103/PhysRevLett.107.040601](https://doi.org/10.1103/PhysRevLett.107.040601).
- [149] B. C. HALL, *An elementary introduction to groups and representations*, arXiv, (2000), [doi:10.48550/arXiv.math-ph/0005032](https://doi.org/10.48550/arXiv.math-ph/0005032).
- [150] J. WURTZ AND P. LOVE, *Maxcut quantum approximate optimization algorithm performance guarantees for $p > 1$* , Phys. Rev. A, 103 (2021), p. 042612, [doi:10.1103/PhysRevA.103.042612](https://doi.org/10.1103/PhysRevA.103.042612).
- [151] M. X. GOEMANS AND D. P. WILLIAMSON, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, 42 (1995), p. 1115–1145, [doi:10.1145/227683.227684](https://doi.org/10.1145/227683.227684).
- [152] E. HALPERIN, D. LIVNAT, AND U. ZWICK, *Max cut in cubic graphs*, Journal of Algorithms, 53 (2004), pp. 169–185, [doi:https://doi.org/10.1016/j.jalgor.2004.06.001](https://doi.org/10.1016/j.jalgor.2004.06.001).
- [153] M. M. WOLF, F. VERSTRAETE, M. B. HASTINGS, AND J. I. CIRAC, *Area laws in quantum systems: Mutual information and correlations*, Phys. Rev. Lett., 100 (2008), p. 070502, [doi:10.1103/PhysRevLett.100.070502](https://doi.org/10.1103/PhysRevLett.100.070502).
- [154] M. KLIESCH AND A. RIERA, *Properties of thermal quantum states: Locality of temperature, decay of correlations, and more*, in *Fundamental Theories of Physics*, Springer International Publishing, 2018, pp. 481–502.
- [155] E. CARLEN, *Trace inequalities and quantum entropy: an introductory course*, Entropy and the Quantum, 529 (2010), pp. 73–140.
- [156] A. IMPARATO, N. CHANCELLOR, AND G. DE CHIARA, *A thermodynamic approach to optimization in complex quantum systems*, Quantum Science and Technology, 9 (2024), p. 025011, [doi:10.1088/2058-9565/ad26b3](https://doi.org/10.1088/2058-9565/ad26b3).

-
- [157] A. W. MARSHALL, I. OLKIN, AND B. C. ARNOLD, *Inequalities: Theory of Majorization and Its Applications*, Springer New York, NY, 2010.
- [158] W. K. HASTINGS, *Monte Carlo sampling methods using Markov chains and their applications*, *Biometrika*, 57 (1970), pp. 97–109, [doi:10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97).
- [159] W. KRAUTH, *Cluster Monte Carlo Algorithms*, John Wiley & Sons, Ltd, 2004, ch. 2, pp. 5–22.
- [160] M. OHKUWA, H. NISHIMORI, AND D. A. LIDAR, *Reverse annealing for the fully connected p -spin model*, *Phys. Rev. A*, 98 (2018), p. 022314, [doi:10.1103/PhysRevA.98.022314](https://doi.org/10.1103/PhysRevA.98.022314).
- [161] Y. YAMASHIRO, M. OHKUWA, H. NISHIMORI, AND D. A. LIDAR, *Dynamics of reverse annealing for the fully connected p -spin model*, *Phys. Rev. A*, 100 (2019), p. 052321, [doi:10.1103/PhysRevA.100.052321](https://doi.org/10.1103/PhysRevA.100.052321).
- [162] M. NIELSEN, *A geometric approach to quantum circuit lower bounds*, *Quantum Information and Computation*, 6 (2006), p. 213–262, [doi:10.26421/qic6.3-2](https://doi.org/10.26421/qic6.3-2).
- [163] M. A. NIELSEN, M. R. DOWLING, M. GU, AND A. C. DOHERTY, *Quantum computation as geometry*, *Science*, 311 (2006), pp. 1133–1135, [doi:10.1126/science.1121541](https://doi.org/10.1126/science.1121541).
- [164] M. DOWLING AND M. NIELSEN, *The geometry of quantum computation*, *Quantum Information and Computation*, 8 (2008), p. 861–899, [doi:10.26421/qic8.10-1](https://doi.org/10.26421/qic8.10-1).
- [165] A. CARLINI, A. HOSOYA, T. KOIKE, AND Y. OKUDAIRA, *Time-optimal quantum evolution*, *Phys. Rev. Lett.*, 96 (2006), p. 060503, [doi:10.1103/PhysRevLett.96.060503](https://doi.org/10.1103/PhysRevLett.96.060503).
- [166] A. CARLINI, A. HOSOYA, T. KOIKE, AND Y. OKUDAIRA, *Time-optimal unitary operations*, *Physical Review A*, 75 (2007), [doi:10.1103/physreva.75.042308](https://doi.org/10.1103/physreva.75.042308).

- [167] A. T. REZAKHANI, W.-J. KUO, A. HAMMA, D. A. LIDAR, AND P. ZANARDI, *Quantum adiabatic brachistochrone*, Physical Review Letters, 103 (2009), [doi:10.1103/physrevlett.103.080502](https://doi.org/10.1103/physrevlett.103.080502).
- [168] X. WANG, M. ALLEGRA, K. JACOBS, S. LLOYD, C. LUPO, AND M. MOHSENI, *Quantum brachistochrone curves as geodesics: Obtaining accurate minimum-time protocols for the control of quantum systems*, Phys. Rev. Lett., 114 (2015), p. 170501, [doi:10.1103/PhysRevLett.114.170501](https://doi.org/10.1103/PhysRevLett.114.170501).
- [169] H. WAKAMURA AND T. KOIKE, *A general formulation of time-optimal quantum control and optimality of singular protocols*, New Journal of Physics, 22 (2020), p. 073010, [doi:10.1088/1367-2630/ab8ab3](https://doi.org/10.1088/1367-2630/ab8ab3).
- [170] D. WANG, H. SHI, AND Y. LAN, *Quantum brachistochrone for multiple qubits*, New Journal of Physics, 23 (2021), p. 083043, [doi:10.1088/1367-2630/ac1df5](https://doi.org/10.1088/1367-2630/ac1df5).
- [171] A. C. SANTOS, C. J. VILLAS-BOAS, AND R. BACHELARD, *Quantum adiabatic brachistochrone for open systems*, Phys. Rev. A, 103 (2021), p. 012206, [doi:10.1103/PhysRevA.103.012206](https://doi.org/10.1103/PhysRevA.103.012206).
- [172] J. YANG AND A. DEL CAMPO, *Minimum-time quantum control and the quantum brachistochrone equation*, arXiv, (2022), [doi:10.48550/ARXIV.2204.12792](https://doi.org/10.48550/ARXIV.2204.12792).
- [173] D. C. BRODY AND D. W. HOOK, *On optimum hamiltonians for state transformations*, Journal of Physics A: Mathematical and General, 39 (2006), pp. L167–L170, [doi:10.1088/0305-4470/39/11/I02](https://doi.org/10.1088/0305-4470/39/11/I02).
- [174] J. ANANDAN AND Y. AHARONOV, *Geometry of quantum evolution*, Phys. Rev. Lett., 65 (1990), pp. 1697–1700, [doi:10.1103/PhysRevLett.65.1697](https://doi.org/10.1103/PhysRevLett.65.1697).
- [175] D. A. FEDOROV, B. PENG, N. GOVIND, AND Y. ALEXEEV, *VQE method: a short survey and recent developments*, Materials Theory, 6 (2022), [doi:10.1186/s41313-021-00032-6](https://doi.org/10.1186/s41313-021-00032-6).

-
- [176] L. LI, M. FAN, M. CORAM, P. RILEY, AND S. LEICHENAUER, *Quantum optimization with a novel Gibbs objective function and ansatz architecture search*, Phys. Rev. Research, 2 (2020), p. 023074, [doi:10.1103/PhysRevResearch.2.023074](https://doi.org/10.1103/PhysRevResearch.2.023074).
- [177] P. K. BARKOUTSOS, G. NANNICINI, A. ROBERT, I. TAVERNELLI, AND S. WOERNER, *Improving variational quantum optimization using CVaR*, Quantum, 4 (2020), p. 256, [doi:10.22331/q-2020-04-20-256](https://doi.org/10.22331/q-2020-04-20-256).
- [178] D. C. BRODY AND D. M. MEIER, *Solution to the quantum Zermelo navigation problem*, Phys. Rev. Lett., 114 (2015), p. 100502, [doi:10.1103/PhysRevLett.114.100502](https://doi.org/10.1103/PhysRevLett.114.100502).
- [179] D. C. BRODY, G. W. GIBBONS, AND D. M. MEIER, *Time-optimal navigation through quantum wind*, New Journal of Physics, 17 (2015), p. 033048, [doi:10.1088/1367-2630/17/3/033048](https://doi.org/10.1088/1367-2630/17/3/033048).
- [180] B. RUSSELL AND S. STEPNEY, *Zermelo navigation and a speed limit to quantum information processing*, Phys. Rev. A, 90 (2014), p. 012303, [doi:10.1103/PhysRevA.90.012303](https://doi.org/10.1103/PhysRevA.90.012303).
- [181] B. RUSSELL AND S. STEPNEY, *Zermelo navigation in the quantum brachistochrone*, Journal of Physics A: Mathematical and Theoretical, 48 (2015), p. 115303, [doi:10.1088/1751-8113/48/11/115303](https://doi.org/10.1088/1751-8113/48/11/115303).
- [182] K. F. RILEY, M. P. HOBSON, AND S. BENCE, *Mathematical Methods for Physics and Engineering - 3rd Edition*, Cambridge University Press, 2006.
- [183] S. BRAVYI AND B. TERHAL, *Complexity of stoquastic frustration-free hamiltonians*, SIAM Journal on Computing, 39 (2010), pp. 1462–1485, [doi:10.1137/08072689X](https://doi.org/10.1137/08072689X).
- [184] A. GALDA, X. LIU, D. LYKOV, Y. ALEXEEV, AND I. SAFRO, *Transferability of optimal QAOA parameters between random graphs*, in 2021 IEEE International Conference on Quantum Computing and Engineering (QCE), 2021, pp. 171–180, [doi:10.1109/QCE52317.2021.00034](https://doi.org/10.1109/QCE52317.2021.00034).

- [185] M. LAPERT, Y. ZHANG, M. BRAUN, S. J. GLASER, AND D. SUGNY, *Singular extremals for the time-optimal control of dissipative spin $\frac{1}{2}$ particles*, Phys. Rev. Lett., 104 (2010), p. 083001, [doi:10.1103/PhysRevLett.104.083001](https://doi.org/10.1103/PhysRevLett.104.083001).
- [186] V. MUKHERJEE, A. CARLINI, A. MARI, T. CANEVA, S. MONTANGERO, T. CALARCO, R. FAZIO, AND V. GIOVANNETTI, *Speeding up and slowing down the relaxation of a qubit by optimal control*, Phys. Rev. A, 88 (2013), p. 062326, [doi:10.1103/PhysRevA.88.062326](https://doi.org/10.1103/PhysRevA.88.062326).
- [187] E. H. LIEB AND D. W. ROBINSON, *The finite group velocity of quantum spin systems*, Communications in Mathematical Physics, 28 (1972), pp. 251–257, [doi:10.1007/BF01645779](https://doi.org/10.1007/BF01645779).
- [188] Z. WANG AND K. R. HAZZARD, *Tightening the Lieb-Robinson bound in locally interacting systems*, PRX Quantum, 1 (2020), p. 010303, [doi:10.1103/PRXQuantum.1.010303](https://doi.org/10.1103/PRXQuantum.1.010303).
- [189] A. M. CHILDS AND N. WIEBE, *Product formulas for exponentials of commutators*, Journal of Mathematical Physics, 54 (2013), p. 062202, [doi:10.1063/1.4811386](https://doi.org/10.1063/1.4811386).
- [190] W. LECHNER, P. HAUKE, AND P. ZOLLER, *A quantum annealing architecture with all-to-all connectivity from local interactions*, Science Advances, 1 (2015), [doi:10.1126/sciadv.1500838](https://doi.org/10.1126/sciadv.1500838).
- [191] P. HAUKE, H. G. KATZGRABER, W. LECHNER, H. NISHIMORI, AND W. D. OLIVER, *Perspectives of quantum annealing: methods and implementations*, Reports on Progress in Physics, 83 (2020), p. 054401, [doi:10.1088/1361-6633/ab85b8](https://doi.org/10.1088/1361-6633/ab85b8).
- [192] A. RAJAK, S. SUZUKI, A. DUTTA, AND B. K. CHAKRABARTI, *Quantum annealing: an overview*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 381 (2022), [doi:10.1098/rsta.2021.0417](https://doi.org/10.1098/rsta.2021.0417).

-
- [193] D. QIAN-HENG, Z. SHUO, W. WEI, AND C. PING-XING, *An alternative approach to construct the initial hamiltonian of the adiabatic quantum computation*, Chinese Physics Letters, 30 (2013), p. 010302, [doi:10.1088/0256-307X/30/1/010302](https://doi.org/10.1088/0256-307X/30/1/010302).
- [194] T. GRASS, *Quantum annealing with longitudinal bias fields*, Phys. Rev. Lett., 123 (2019), p. 120501, [doi:10.1103/PhysRevLett.123.120501](https://doi.org/10.1103/PhysRevLett.123.120501).
- [195] K. BOOTHBY, P. BUNYK, J. RAYMOND, AND A. ROY, *Next-generation topology of d-wave quantum processors*, arXiv, (2020), [doi:10.48550/arXiv.2003.00133](https://doi.org/10.48550/arXiv.2003.00133).
- [196] K. BOOTHBY, C. ENDERUD, AND T. LANTING ET AL., *Architectural considerations in the design of a third-generation superconducting quantum annealing processor*, arXiv, (2021), [doi:10.48550/arXiv.2108.02322](https://doi.org/10.48550/arXiv.2108.02322).
- [197] A. SELBY, *Efficient subgraph-based sampling of ising-type models with frustration*, arXiv, (2014), [doi:10.48550/arXiv.1409.3934](https://doi.org/10.48550/arXiv.1409.3934).
- [198] R. S. ANDRIST, M. J. A. SCHUETZ, AND P. MINSEN ET AL., *Hardness of the maximum-independent-set problem on unit-disk graphs and prospects for quantum speedups*, Phys. Rev. Res., 5 (2023), p. 043277, [doi:10.1103/PhysRevResearch.5.043277](https://doi.org/10.1103/PhysRevResearch.5.043277).
- [199] A. PALACIOS, A. GARCIA-SAEZ, AND M. P. ESTARELLAS, *A scalable 2-local architecture for quantum annealing of ising models with arbitrary dimensions*, arXiv, (2024), [doi:10.48550/arXiv.2404.06861](https://doi.org/10.48550/arXiv.2404.06861).